

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries

PT as central nervous system injuries.

```
XX
PS Example 5; SEQ ID NO 3368; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA038642-AA042213) with nucleotidic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 229 AA;
XX
Query Match 100.0%; Score 1198; DB 4; Length 229;
Best Local Similarity 100.0%; Pred. No. 1e-126;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MAAPPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEMDTQVVKSSPLGPAGLGAEEP 60
1 MAAPPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEMDTQVVKSSPLGPAGLGAEEP 60
DB 1 MAAPPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEMDTQVVKSSPLGPAGLGAEEP 60
XX
QY 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120
61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120
DB 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120
XX
QY 121 VGIKESLKGSTYNLLFCGSGGIPVGFHLVSTHAALALRGHFCLSDDKXVCYLKTKAY 180
121 VGIKESLKGSTYNLLFCGSGGIPVGFHLVSTHAALALRGHFCLSDDKXVCYLKTKAY 180
DB 121 VGIKESLKGSTYNLLFCGSGGIPVGFHLVSTHAALALRGHFCLSDDKXVCYLKTKAY 180
XX
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTNHNLSLMLKLTSEVTPQSKPEN 229
181 NASEMDIQNVPLSEKIAELKEKIVLTNHNLSLMLKLTSEVTPQSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTNHNLSLMLKLTSEVTPQSKPEN 229
XX
RESULT 2
ABG34856
ID ABG34856 strand; protein; 229 AA.
XX
AC ABG34856;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human cancer related protein encoded by cDNA 85PIB3.
XX
KM Human; cytostatic; 85PIB3; cancer; immunogen; chromosome 15q14.
XX
OS Homo sapiens.
XX
PN WO200218578-A2.
XX
PD 07-MAR-2002.
XX
PF 28-AUG-2001; 2001WO-US026838.
XX
PR 28-AUG-2000; 2000US-0228432P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Raitano AB, Farris W, Hubert RS, Afar D, Ge W, Challita-Eid P;
PI Jakobovits A;
XX
-XX WPI: 2002-382963/41.
DR N-PSDB; ABR70506.
XX
PT Composition for modulating the status of 85PIB3 protein or a molecule
```

```
PT comprising a substance e.g. antibody specific to, nucleic acid encoding,
PT or ribozyme of 85PIB3.
XX
PS Claim 34; Fig 2; 201pp; English.
XX
CC The invention relates to a composition comprising a substance that
CC modulate the status of 85PIB3, where the status of a cell expresses
CC 85PIB3 gene product is modulated. Also included are a composition
CC comprising a peptide region of 5 amino acids of the 85PIB3 protein, in
CC any whole number increment up to 229 that includes an aa position
CC selected from an aa position having a value greater than 0.5 in the
CC hydrophilicity profile, an aa position having a value less than 0.5 in
CC the hydrophobicity profile, an aa position having a value greater than
CC 0.5 in the percent accessible residue profile, an aa position having a
CC value greater than 0.5 in the average flexibility profile, or an aa
CC position having a value greater than 0.5 in the beta-turn profile; a
CC polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous
CC residues of the 85PIB3 protein; a recombinant protein comprising the
CC antigen-binding region of a monoclonal antibody; a non-human transgenic
CC animal that produces an antibody that binds to the 85PIB3 protein; a
CC hybridoma that produces antibody specific to the protein; a single chain
CC monoclonal antibody (MAb) that comprises the variable domains of the
CC heavy and monoclonal antibodies specific to the protein; a vector
CC comprising a polynucleotide that encodes the MAb; inhibiting growth of
CC cancer cells or treating a patient who bears cancer cells that expresses
CC the protein, by administering the protein, antibody, polynucleotide
CC encoding the protein, antisense polynucleotide to the polynucleotide,
CC ribozyme that cleaves the polynucleotide and T cells that specifically
CC recognize the protein, and generating a mammalian immune response
CC directed to the protein exposing cells of the mammal's immune system to
CC an immunogenic portion of the protein or polynucleotide. The composition,
CC which comprises an antibody specific to the protein, is useful for
CC delivering a cytotoxic agent to a cell that expresses the protein by
CC providing a cytotoxic agent conjugated to antibody and exposing the cell
CC to the antibody-agent conjugate. The methods are useful for inhibiting
CC growth of cancer cells or treating a patient who bears cancer cells that
CC expresses the protein, for generating a mammalian immune response
CC directed to the protein, for detecting the presence of the protein or
CC polynucleotide in a biological sample in a patient who has or who is
CC suspected of having cancer and for monitoring 85PIB3 in a biological
CC sample from a patient who has or who is suspected of having cancer. The
CC gene for 85PIB3 is located on human chromosome 15q14. The present
CC sequence is the 85PIB3 protein
XX
SQ Sequence 229 AA;
XX
Query Match 100.0%; Score 1198; DB 5; Length 229;
Best Local Similarity 100.0%; Pred. No. 1e-126;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MAAPPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEMDTQVVKSSPLGPAGLGAEEP 60
1 MAAPPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEMDTQVVKSSPLGPAGLGAEEP 60
DB 1 MAAPPLRHRSRCATPPRGDFCGGTERAIDQASFTTSMEMDTQVVKSSPLGPAGLGAEEP 60
XX
QY 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120
61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120
DB 61 AAGPOLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSRSLGAVFSRVTNNVLEAPFL 120
XX
QY 121 VGIKESLKGSTYNLLFCGSGGIPVGFHLVSTHAALALRGHFCLSDDKXVCYLKTKAY 180
121 VGIKESLKGSTYNLLFCGSGGIPVGFHLVSTHAALALRGHFCLSDDKXVCYLKTKAY 180
DB 121 VGIKESLKGSTYNLLFCGSGGIPVGFHLVSTHAALALRGHFCLSDDKXVCYLKTKAY 180
XX
QY 181 NASEMDIQNVPLSEKIAELKEKIVLTNHNLSLMLKLTSEVTPQSKPEN 229
181 NASEMDIQNVPLSEKIAELKEKIVLTNHNLSLMLKLTSEVTPQSKPEN 229
DB 181 NASEMDIQNVPLSEKIAELKEKIVLTNHNLSLMLKLTSEVTPQSKPEN 229
XX
RESULT 3
AAM42009
ID AAM42009 strand; protein; 231 AA.
XX
AC AAM42009;
XX
XX
```

BT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6940.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552312.
 PR 20-JUN-2000; 2000US-00596042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AA161165.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 2; SEQ ID NO 6940; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA38642-AA42213) with nootropic, and the
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 231 AA;
 XX
 Query Match 100.0%; Score 1198; DB 4; Length 231;
 Best Local Similarity 100.0%; Pred. No. 1e-126;
 Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAAGPRLRRSRCAATPRGDFCGGTERAIDQASFTTSMEMDTQYVKSSPLGPAAGABEP 60
 DB 3 MAAGPRLRRSRCAATPRGDFCGGTERAIDQASFTTSMEMDTQYVKSSPLGPAAGABEP 62
 QY 61 AAGPOLPSWLOPBRCAVFCQACQAVLAADSVHLAWLSSLSGAVVSRTYNNVLEAPPL 120
 DB 63 AAGPOLPSWLOPBRCAVFCQACQACQAVLAADSVHLAWLSSLSGAVVSRTYNNVLEAPPL 122

QY 121 VGIEGSLKGSYVNLFFCGSGGIPVGFPHLYSTHAALALRGHFLCLSDRWVYLLKTKATV 180
 DB 123 VGIEGSLKGSYVNLFFCGSGGIPVGFPHLYSTHAALALRGHFLCLSDRWVYLLKTKATV 182
 QY 181 NASEMDIQNVPLSEKTAELKEKIVLTHNRUKSLMKLISEVTPQOSKRN 229
 DB 183 NASEMDIQNVPLSEKTAELKEKIVLTHNRUKSLMKLISEVTPQOSKRN 231
 RESULT 4
 ABG34855
 ID ABG34855 standard; protein; 164 AA.
 XX
 AC ABG34855;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cDNA 85P1B3 splice variant, open reading frame #3.
 XX
 DE Human; cytostatic; 85P1B3; cancer; immunogen; chromosome 15q14.
 XX
 OS Homo sapiens.
 XX
 PN WO200218578-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 28-AUG-2001; 2001WO-US026838.
 XX
 PR 28-AUG-2000; 2000US-0228432P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Raitano AB, Farris M, Hubert RS, Afar D, Ge W, Challita-Bid P,
 PI Jakovovits A;
 XX
 DR WPI; 2002-382963/41.
 DR N-PSDB; ABK70504.
 XX
 PT Composition for modulating the status of 85P1B3 protein or a molecule
 PT comprising a substance e.g. antibody specific to, nucleic acid encoding,
 PT or ribozyme of 85P1B3.
 XX
 PS Example 38; Page 124; 201pp; English.
 XX
 CC The invention relates to a composition comprising a substance that
 CC modulate the status of 85P1B3, where the status of a cell expresses
 CC 85P1B3 gene product is modulated. Also included are a composition
 CC comprising a peptide region of 5 amino acids of the 85P1B3 protein, in
 CC any whole number increment up to 229 that includes an aa position
 CC selected from an aa position having a value greater than 0.5 in the
 CC hydrophilicity profile, an aa position having a value less than 0.5 in
 CC the hydrophobicity profile, an aa position having a value greater than
 CC 0.5 in the percent accessible residue profile, an aa position having a
 CC value greater than 0.5 in the average flexibility profile, or an aa
 CC position having a value greater than 0.5 in the beta-turn profile; a
 CC polynucleotide that encodes analogue peptide of 8, 9, 10 or 11 contiguous
 CC residues of the 85P1B3 protein; a recombinant protein comprising the
 CC antigen-binding region of a monoclonal antibody; a non-human transgenic
 CC animal that produces an antibody that binds to the 85P1B3 protein; a
 CC hybridoma that produces an antibody specific to the protein; a single chain
 CC monoclonal antibody (MAb) that comprises the variable domains of the
 CC heavy and monoclonal antibodies specific to the protein; a vector
 CC comprising a polynucleotide that encodes the MAb; inhibiting growth of
 CC cancer cells or treating a patient who bears cancer cells that expresses
 CC the protein, by administering the protein, antibody, polynucleotide
 CC encoding the protein, antisense polynucleotide to the polynucleotide,
 CC ribozyme that cleaves the polynucleotide and T cells that specifically
 CC recognize the protein, and generating a mammalian immune response
 CC directed to the protein exposing cells of the mammal's immune system to
 CC an immunogenic portion of the protein or polynucleotide. The composition,
 CC which comprises an antibody specific to the protein, is useful for
 CC delivering a cytotoxic agent to a cell that expresses the protein by

	Query Match	9.4%	Score 113;	DB 4;	Length 233;
	Beet Local Similarity	27.1%;	Pred. No. 0.0012;		
	Matches	52;	Conservative	30;	No. matches 88; Indels 22; Gaps 8
Qy	31 ASFTTSMEMDQQVVKGSPLCPAGIGAEPPAGEPOLPSMLQPERCARVEQCACQACHAVLADS	90			
	: : : : :				
Dd	49 ASMMSHSESDSV---ADMERAPQL--EEEAAMAE-----ERLYVELCSCCRPFLQDS	95			
Qy	91 VHLAMDLSR-SLGAVVFSPRTVNIVLEAPFLVGIEGSLKSTYNNLLFCGSGCGI PVGFILY	149			
	: : : : : : : : : : : : : :				
Dd	96 --LFWVASGEPTNCLILRCVCSNVSDVDQGLSKREKENGCYLERTLCAGGSLNNGYYR	153			

Qy 150 STAAALALRGHCSSDDGVCYL--KTKIVANSEMDIONVPLSEKIALTKKIVLTH 207
 Db 154 CTRKNDYKRDPLFCLSVENIESVIGSEKQIV-SEKELTNL---ESRVEIEKSLQME 209
 Qy 208 NRKLSKMLTISE 219
 Db 210 DVKFKALQMKLME 221

CC	RESULT 6
XX	AAG93290
ID	AAG93290 standard; protein; 233 AA.
XX	
AC	AAG93290;
XX	
DT	13-SEP-2001 (first entry)
XX	
DE	Human protein HP10650.
XX	
KW	Human; gene therapy; tumour.
XX	
OS	Homo sapiens.
XX	
PN	WO200142302-A1.
XX	
PD	14-JUN-2001.
XX	
PF	06-DEC-2000; 2000WO-JP008631.
XX	
PR	06-DEC-1999; 99JP-00346863.
XX	
PR	06-DEC-1999; 99JP-00346864.
XX	
PR	08-FEB-2000; 2000JP-00031062.
XX	
PR	10-FEB-2000; 2000JP-00034090.
XX	
PR	10-FEB-2000; 2000JP-00034091.
XX	
PR	14-FEB-2000; 2000JP-00035829.
XX	
PR	14-FEB-2000; 2000JP-00035899.
XX	
PR	14-MAR-2000; 2000JP-00071161.
XX	
PR	30-MAY-2000; 2000JP-00160851.
XX	
PA	(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX	
PI	Kato S, Eguchi C, Saeki M;
XX	
DR	WPI; 2001-381646/40.
XX	
DR	N-PSDB; AAH68575.
XX	
PT	Human protein originated from tumor cell line, applicable as drug,
XX	
PT	reagent for studying intracellular protein networks and protein source
XX	
PT	for drug screening, also encoded cDNA for gene diagnosis and gene
XX	
PT	therapy.
XX	
PS	Claim 1; Page 303-304; 471pp; Japanese.
XX	
XX	
CC	The present sequence is a human protein. The human protein, preferably
XX	
CC	originated from tumour cell line, is applicable as a drug, a reagent for
XX	
CC	studying intracellular protein networks and a protein source for
XX	
CC	screening proteins for binding low molecular weight drugs. The human
XX	
CC	protein coding sequence is useful for gene diagnosis and gene therapy,
XX	
CC	expression vectors and transformant cells for detection of ligands and
XX	
CC	receptors

```

XX      SQ      Sequence 233 AA;
CC      Query Match          9.4%; Score 113; DB 4; Length 233;
CC      Best Local Similarity 27.1%; Pred. No. 0.0012;
CC      Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY      31 ASFTSMEMDTQVVKSSPLGPAGLGAEPAPGQPLPSWIQPERCAVFOCAQCHAVLADS 90
DB      49 ASMWSSMSSEDAV---ADMERAPQ--EEERAAAE-----ERPLVFLCSGCRRLPGDS 95
QY      91 VHLAMDLSR-SLGAVFSRYTNVNVLEAPLVGIEGSLKSTYVNLFCGSGCIGPVGFHLY 149
DB      96 --LSWVASQEDTNCILRCVSCNVSVDEKQKLSREKENGCVLETLCCAGCSINLGYVYR 153
QY      150 STHAALALRGHFLCSDDKRVCYLL--KTAIVNASMDIQVPLSEKIAELKEKIVLTH 207
DB      154 CTRPNLDYKRDLCFLSVEAIESYVLGSSSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209
QY      208 NRLKSLMKILSE 219
DB      210 DVLKALQMKLWE 221

RESULT 7
ADJ69163
ID      ADJ69163 standard; protein; 233 AA.
XX      AC      ADJ69163;
XX      DT      06-MAY-2004 (first entry)
XX      DE      Human heat mitochondrial protein as a therapeutic target SegID969.
XX      KW      mitochondrial; human; screening assay; diabetes mellitus;
XX      KW      Huntington's disease; osteoarthritis;
XX      KW      Leber's hereditary optic neuropathy; LHON;
XX      KW      mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX      KW      myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX      KW      neuroprotective; noctropic; antidiabetic; anticonvulsant; antiarthritic;
XX      KW      osteopathic; ophthalmological; cytostatic.
XX      OS      Homo sapiens.
XX      PN      MO200308766-A2.
XX      PD      23-OCT-2003.
XX      PF      04-APR-2003; 2003MO-US010870.
XX      PR      12-APR-2002; 2002US-0372843P.
XX      PR      17-JUN-2002; 2002US-0389987P.
XX      PR      20-SEP-2002; 2002US-0412418P.
XX      PA      (MITO-) MITOKOR.
XX      PA      (BUCK-) BUCK INST AGE RES.
XX      PI      Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
XX      PI      Warnock DE;
XX      DR      MPI; 2003-845369/78.
XX      PT      Identifying a mitochondrial target for drug screening assays and for
XX      PT      treating diseases associated with altered mitochondrial function,
XX      PT      comprises detecting a modified polypeptide in a sample and correlating
XX      PT      with the disease.
XX      PS      Claim 1; SEQ ID NO 969; 180pp; English.
XX      CC      This invention relates to novel mitochondrial targets that can be used
XX      CC      for therapeutic intervention in treating a disease associated with
XX      CC      altered mitochondrial function. Specifically, it refers to a method for
XX      CC      identifying proteins of the human heart mitochondrial proteome that are

```

```

CC      useful for drug screening assays, as well as therapeutic targets. The
CC      present invention describes a method for identifying such proteins that
CC      can be used in the treatment of various diseases associated with altered
CC      mitochondrial function including diabetes mellitus, Huntington's disease,
CC      osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC      encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC      ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC      compositions have neuroprotective, noctropic, antidiabetic,
CC      anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC      cyostatic activities. This polypeptide sequence is a human heart
CC      mitochondrial protein of the invention.
XX      SQ      Sequence 233 AA;
CC      Query Match          9.4%; Score 113; DB 7; Length 233;
CC      Best Local Similarity 27.1%; Pred. No. 0.0012;
CC      Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY      31 ASFTSMEMDTQVVKSSPLGPAGLGAEPAPGQPLPSWIQPERCAVFOCAQCHAVLADS 90
DB      49 ASMWSSMSSEDAV---ADMERAPQ--EEERAAAE-----ERPLVFLCSGCRRLPGDS 95
QY      91 VHLAMDLSR-SLGAVFSRYTNVNVLEAPLVGIEGSLKSTYVNLFCGSGCIGPVGFHLY 149
DB      96 --LSWVASQEDTNCILRCVSCNVSVDEKQKLSREKENGCVLETLCCAGCSINLGYVYR 153
QY      150 STHAALALRGHFLCSDDKRVCYLL--KTAIVNASMDIQVPLSEKIAELKEKIVLTH 207
DB      154 CTRPNLDYKRDLCFLSVEAIESYVLGSSSEKQIV-SEDKELFNL---ESRVEIEKSLTQME 209
QY      208 NRLKSLMKILSE 219
DB      210 DVLKALQMKLWE 221

RESULT 8
ADJ31517
ID      ADJ31517 standard; protein; 233 AA.
XX      AC      ADJ31517;
XX      DT      20-MAY-2004 (first entry)
XX      DE      Human protein encoded by a full length cDNA clone SegID 3550.
XX      KW      human; medicine; signal transduction; glycoprotein; transcription;
XX      KW      oligo-capping method.
XX      OS      Homo sapiens.
XX      PN      EP1396543-A2.
XX      PD      10-MAR-2004.
XX      PF      07-JUL-2000; 2003EP-00025638.
XX      PR      08-JUL-1999; 99JP-00194486.
XX      PR      11-JAN-2000; 2000JP-0018774.
XX      PR      02-MAY-2000; 2000JP-00183865.
XX      PR      07-JUL-2000; 2000EP-00114089.
XX      PA      (REAS-) RES ASSOC BIOTECHNOLOGY.
XX      PI      Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
XX      PI      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX      DR      MPI; 2004-204755/20.
XX      DR      N-PSDB; ADJ31516.
XX      PT      New oligonucleotide primers (830 cDNAs) useful for synthesizing full
XX      PT      length human cDNAs.
XX      PS      Example 1; SEQ ID NO 3550; 1340pp; English.

```


CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 695 AA;

Query Match 7.7%; Score 92; DB 8; Length 695;
 Best Local Similarity 25.3%; Pred. No. 1.5; Mismatches 38; Conservative 28; Indels 34; Gaps 7;

QY 101 LGAVFSRYTNVNLBAPFLVGIKSTYMLFCGSCGIPVGHLYSTHAALALRG 160
 DB 101 LGITIR-----FVSGTFFPSGAKGELSKRKPMMMLITWGITVAV AVSVATIMSING 153
 QY 161 HF-----CLSSDKMVCYLKTKAIVNASE-MDI-----QNVPLSE-K 195
 DB 154 HMGGMFWFELATLIVMLIGHLIEMKAIMGAGDALXDLASTLVPRKAKLKSQDELSELK 213
 QY 196 IAEIKERKIVLTHNRKSLMKIKLSEVTPPOS 225
 DB 214 VGDL--LIVKENKIPADGILISEALVDES 241

RESULT 11

AAO08972
 ID AAO08972 standard; protein; 269 AA.

AC AAO08972;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 22864.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PP 26-FEB-2001; 2001WO-US004927.

PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSB-) HYSBQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR N-PSDB; AAI88903.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 22864; 139pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 269 AA;

Query Match 7.5%; Score 90; DB 4; Length 269;
 Best Local Similarity 23.0%; Pred. No. 0.61; Mismatches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HSRRCATPPRGDFC-----GTERAIDQA-----SFTSMEMDTQVKKSSPL----- 50
 DB 39 YRRQAQPPH---CPABEGEPGAPQALGAPSTSVSLTAVAVDPVCPLOGSHALCTCCFQ 95
 QY 51 ---GPAGLGAEPPAAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSR----- 99
 DB 96 PMPDRPAERBQDPRVAPQ-----QCAVC---LQPCHLIYWGCTRGCGCLA 139
 QY 100 -----SLGAVFSRYTNVNLBAPFLVGIKSTY-NLFCGSCGIPVGHLYSTHA 153
 DB 140 PFCGLNLGPKCUDGVANNNSYSDILKNTLAT-RGLTWKML-----TES 183
 QY 154 ALAALRGHCLG-----SDKMCYLLKTKAIVNASEMDIONVPLSE 194
 DB 184 LMALQRGVFLSDRYVTGDTVLCCYCCGLRSFRELTYQYRONIPASE 229

RESULT 12

ABM83354
 ID ABM83354 standard; protein; 572 AA.

AC ABM83354;

DT 18-NOV-2004 (first entry)

XX Human diagnostic and therapeutic protein SEQ ID NO:3603.

KW Gene therapy; human diagnostic and therapeutic polynucleotide; dtbp.

KW Homo sapiens.

PN WO2004023973-A2.

PD 25-MAR-2004.

PP 12-SEP-2003; 2003WO-US028227.

PR 12-SEP-2002; 2002US-0410259P.

PR 12-SEP-2002; 2002US-0410260P.

XX (INCY-) INCYTE CORP.

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,

PI Harthorne TA, Suchanowski MT, Altus CM, Plets SJ, Elder LV;

PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;

PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;

PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Mingrove J, Vilt UA, Kitson ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patuxy S, Shi X, Suarez CJ;
 DR WPI; 2004-329368/30.
 DR N-PSDB; ACN42006.
 XX
 PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 PS Claim 27; Page; 190pp; English.
 CC
 CC The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders, endocrine
 CC autoimmune/inflammatory disorder, developmental disorder, or
 CC disorder, neurological disorder, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dthp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dthp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/11listing.htm
 CC
 SQ Sequence 572 AA;

Query Match 7.4%; Score 89; DB 8; Length 572;
 Best Local Similarity 23.0%; Pred. No. 2.5;
 Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
 QY 8 HRSRCATPRRGDRC-----GTERAIDQ-----SFTSMENDTVVKGSSPL----- 50
 DB 342 YRROAOPPH---CPAPEGEPGAPQALGDAPSTSVSLTTAVQDYVCPDQSHALCTCCFQ 398
 QY 51 ---GPAGLGAEBPAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSR----- 99
 DB 399 PMEDRAERERQDRVAQ-----QCAVC---LQPCHLWYGCTRRGCVGCLA 442
 QY 100 ----SIGAVVFSRVTNVNVLEAPLVIGISLKGSTY-NLFCGSGCIPVGFHLVSTHA 153
 DB 443 PFCELNLGDKCLDGVNLNNSYESDILKNVLYAT-RGLTWKMKL-----TES 486
 QY 154 ALALRGHFCLS-----SDKMVCYLKTKAIVNASEMDIONVPLSE 194
 DB 487 LVALLRGVFLSLDYRTVGTDLVLCYCGLRSPRELTYYQONITPASE 532

RESULT 13
 AAB93182
 ID AAB93182 standard; protein; 623 AA.
 XX
 AC AAB93182;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 -DE Human protein sequence SEQ ID NO:12128.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 FX
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX

PF 28-JUL-2000; 2000EP-00116126.
 XX
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00118767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12128; 2537bp + Sequence Listing; English.

CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification; where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 XX
 SQ Sequence 623 AA;

Query Match 7.4%; Score 89; DB 4; Length 623;
 Best Local Similarity 23.0%; Pred. No. 2.9;
 Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
 QY 8 HRSRCATPRRGDRC-----GTERAIDQ-----SFTSMENDTVVKGSSPL----- 50
 DB 393 YRROAOPPH---CPAPEGEPGAPQALGDAPSTSVSLTTAVQDYVCPDQSHALCTCCFQ 449
 QY 51 ---GPAGLGAEBPAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSR----- 99
 DB 450 PMEDRAERERQDRVAQ-----QCAVC---LQPCHLWYGCTRRGCVGCLA 493
 QY 100 ----SIGAVVFSRVTNVNVLEAPLVIGISLKGSTY-NLFCGSGCIPVGFHLVSTHA 153
 DB 494 PFCELNLGDKCLDGVNLNNSYESDILKNVLYAT-RGLTWKMKL-----TES 537
 QY 154 ALALRGHFCLS-----SDKMVCYLKTKAIVNASEMDIONVPLSE 194
 DB 538 LVALLRGVFLSLDYRTVGTDLVLCYCGLRSPRELTYYQONITPASE 583

RESULT 14
 ABB97233
 ID ABB97233 standard; protein; 623 AA.
 XX
 AC ABB97233;
 AC


```

XX 27-JUN-2002 (first entry)
DT XX
XX Novel human protein SEQ ID NO: 501.
DE XX
XX Human; antinaemic; vulnerary; antiinflammatory; immunomodulator;
XX antiinfectivity; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX
XX Homo sapiens.
OS
XX WO200222660-A2.
XX
XX 21-MAR-2002.
XX
XX 10-SEP-2001; 2001WO-US026015.
XX
XX 11-SEP-2000; 2000US-00659671.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F,
XX Xue AJ, Yang Y, Wehtman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
XX
XX N-PSDB; ABN32419.
XX
XX Example 2; SEQ ID NO 501; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a protein of the invention
XX
XX Sequence 623 AA;
SO
Query Match 7.4%; Score 89; DB 5; Length 623;
Best Local Similarity 23.0%; Pred. No. 2.9;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
QY 8 HSRCATPPRGDPC-----GTERAIDQA-----SFTSMEMDTGVKGSPL----- 50
DB 393 YRRQAQPPH---CPAPEGEPGAPQALGDAPPTSVSLTAVADYVCPQLQSHALCTCCPQ 449
QY 51 ---GPAGLGAEPGAPGQPLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
DB 450 PMPDRRAERQDPRVAPQ-----QCAVC---LQPCILYMGCTRTGCGCGLA 493
QY 100 ----SLGAVVFSRVNNTNVLEAPFLVGIKSGSTY-NLFCGSCGIPVGFHLYSTHA 153
DB 494 PFCELNLDGKCLDGVNINNSYESDILKNYLAT-RGLTWKML-----TTS 537
QY 154 ALAAGHPLS-----SPKMVYLKTKRAIVNASEMDIONVPLSE 194
DB 538 LVNLQGVFLSLSDYRVVTGDTVLCYCCGSLSFRELTVOYQONIPASE 583

```

```

DT 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12100.
DE XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
OS
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX
XX 27-AUG-1999; 99JP-00300253.
XX
XX 11-JAN-2000; 2000JP-00118776.
XX
XX 02-MAY-2000; 2000JP-00183767.
XX
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Oka T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
XX Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesising polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesising polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX Sequence 652 AA;
SO
Query Match 7.4%; Score 89; DB 4; Length 652;
Best Local Similarity 23.0%; Pred. No. 3.1;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
QY 8 HSRCATPPRGDPC-----GTERAIDQA-----SFTSMEMDTGVKGSPL----- 50
DB 422 YRRQAQPPH---CPAPEGEPGAPQALGDAPSTSVSLTAVADYVCPQLQSHALCTCCPQ 478
QY 51 ---GPAGLGAEPGAPGQPLPSWLQPERCAVFOCAQCHAVLADSVHLAWDLR----- 99
DB 479 PMPDRRAERQDPRVAPQ-----QCAVC---LQPCILYMGCTRTGCGCGLA 522
QY 100 ----SLGAVVFSRVNNTNVLEAPFLVGIKSGSTY-NLFCGSCGIPVGFHLYSTHA 153

```

Db 523 PFCENLGDKCLDGVLNNSYESDILKNYLAT-RGLTWKML-----TES 566

Oy 154 ALAALRGHPCLG-----SDRMVCIYLKTKAIYNASENDIONVPLSE 194
:| | | | | :| | | | | :| | | | |
Db 567 LVVALORGVPILSDYRVYTGDTVLCCGCLRSFRELTVOYRONIPASE 612

Search completed: May 18, 2006, 15:35:47
Job time : 197 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2006, 15:36:06 ; Search time 39 Seconds

(without alignments)
564.965 Million cell updates/sec

Title: US-09-942-052a-728

Perfect score: 1198

Sequence: 1 MAOGLRHRRCATPPRGDF.....LKSIMKILSEVTPDQSPEN 229

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	7.7	695	2 C86731	copper-potassium t
2	90.5	7.6	597	2 D71293	probable phosphori
3	89	7.4	306	2 T46399	hypothetical prote
4	88	7.3	250	2 A43623	kanamycin kinase (
5	86.5	7.2	669	2 A97229	NAD-dependent DNA
6	85	7.1	2194	1 J01977	glutamate synthase
7	84.5	7.1	361	2 AE2074	ferrichrome bindin
8	83	6.9	663	2 D97047	DNA ligase (NAD de
9	82.5	6.9	143	2 B69099	sugar transporter
10	82.5	6.9	493	2 A85433	formate hydrogenu
11	82	6.8	967	2 F87678	DNA polymerase I (
12	81.5	6.8	329	2 D90404	transport protein,
13	81.5	6.8	457	2 H85767	glucuronide permea
14	81.5	6.8	457	2 C90919	glucuronide permea
15	81.5	6.8	457	2 B64918	glucuronide permea
16	81.5	6.8	640	2 T41977	hypothetical prote
17	81.5	6.8	662	2 T44221	probable capsid pr
18	81	6.8	1418	2 S40764	hypothetical prote
19	80.5	6.7	766	2 B85440	receptor kinase-11
20	80.5	6.7	1711	1 A55148	protein-tyrosine-p
21	80	6.7	469	2 S55167	IME2-dependent sig
22	79	6.6	386	2 H90789	probable aminometh
23	79	6.6	386	2 C85650	probable aminometh
24	79	6.6	437	2 C86823	GMP-binding protei
25	78.5	6.6	188	2 F83816	late competence op
26	78.5	6.6	578	2 F86484	probable hydroxyme
27	78.5	6.6	697	1 Q08BK3	UL104 protein - hu
28	78	6.5	351	2 H82098	UDP-3-O-3-hydroxym
29	77.5	6.5	662	2 T44036	hypothetical prote

30	77.5	6.5	1013	2 T31211	trwC protein homol
31	77.5	6.5	1643	2 T14274	versican precursor
32	77	6.4	381	2 AD1113	hypothetical prote
33	77	6.4	721	2 F87611	TonB-dependent rec
34	77	6.4	738	1 TFDHM	melanotransferrin
35	76.5	6.4	307	2 AG2017	glycerol-3-phospha
36	76.5	6.4	387	2 D69392	probable acyl-CoA
37	76.5	6.4	388	2 S57526	cellulase - Fibrob
38	76.5	6.4	764	2 H98143	cbbeC protein (U60
39	76.5	6.4	764	2 AD3144	formate dehydrogen
40	76.5	6.4	1554	2 T06370	probable DNA (cyto
41	76	6.3	367	2 AH0936	glycerol dehydroge
42	75.5	6.3	310	2 A55053	endothelial monocy
43	75.5	6.3	415	2 T46716	hypothetical prote
44	75.5	6.3	700	2 D70951	probable UvrD - My
45	75.5	6.3	1534	2 S59604	DNA (cytosine-5-)-

ALIGNMENTS

RESULT 1

C86731 copper-potassium transporting ATPase B cobP [imported] - Lactococcus lactis subsp. lact

C/Species: Lactococcus lactis subsp. lactis

C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C/Accession: C86731

R/Bolotin, A.; Mincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weisenbach, J.; Ehrh

Genome Res. 11, 731-753, 2001

A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s

A/Reference number: AB6625; NCBI:21235186; PMID:1137471

A/Accession: C86731

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-695 <STO>

A/Cross-references: UNIPROT:Q9CH87; UNIPARC:UPI000006919; GB:AE005176; PID:G12723778;

A/Experimental source: strain IL1403

C/Genetics:

A/Genes: cobP

C/Superfamily: Enterococcus copper-transporting ATPase cobP; ATPase nucleotide-binding

Query Match	7.7%; Score 92; DB 2; Length 695;
Best Local Similarity	25.3%; Pred. No. 2;
Matches	38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
QY	101 LGAVFSRYTNVLEAPLVGIEGSLKSTYNLFCGSCGIPVGFHLYSTHAALARG 160
DB	101 LGTIIF-----FYSGTPFSGAKGELKSRKPAMMLITMGITVAV-AVSVAITMSLNG 153
QY	161 HF-----CISSDKMCYLLKTKATVNASF-MDI-----QNVPLSE-K 195
DB	154 HNGMNFWEFLATLIVMLIGHLIEMKALMGAGDALCVLPKKAHLKSGKVELSELK 213
QY	196 IAELEKEIVLTNRRLKSLMKILSEVTPDOS 225
DB	214 VGDL--LIVKENKIPADGLIILSEALVDES 241
RESULT 2	
D71293	probable phosphoribosylglycinamide formyltransferase - syphilis spirochete
C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)	
C/Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004	
C/Accession: D71293	
R/Frazer, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwi	
lson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uteerback, T.; Mcd	
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.	
Science 281, 375-388, 1998	
A/Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.	
A/Reference number: A71250; NCBI:98332770; PMID:9665876	
A/Accession: D71293	
A/Status: preliminary; nucleic acid sequence not shown; translation not shown	
A/Molecule type: DNA	

A:Residues: 1-597 <COL>
A:Cross-references: UNIPROT:O63693; UNIPARC:UPI0000003290; GB:AE001243; GB:AE000520; NID
A:Experimental source: strain Nicholas
C:Genetic:
A:Gene: TP0695

Query Match
Best Local Similarity 28.0%; Pred. No. 2.2;
Matches 51; Conservative 19; Mismatches 47; Indels 65; Gaps 12;

QY 12 CATPRGDFCGCTERAID-----QASFT-----TSMEMDTQVKGSSPLGPA 53
DB 90 CALP--GHRLEATKATADKTRMRACFTRARLRCPRTFLEPDSFAMDT-----PPGHA 140
QY 54 GLGAEBAAGPQLPQSWLQP-ERCAVFOC--AQCHAVLADSVHLAMDLSRLGAVPSRVT 110
DB 141 RLCSHLSAGLSPELVVVKPTDMMGARGCTLAQCDDTLINACAVARQFSRS----- 190
QY 111 NNVLLEAPFLVGEISLKGSTYNLLFCGSCGIPVGFHLYSTHAALALRGHFCISDKMV 170
DB 191 GRVITE-EPVIGREPSLEG---LIPDGT-----LYVT--ALA-----DRHI 225
QY 171 CY 172
DB 226 CF 227

RESULT 3
T46399
hypothetical protein DKFZp434N2420.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46399
R:Oetemaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23031
A:Accession: T46399
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-306 <AAA>
A:Cross-references: UNIPROT:Q96EP1; UNIPARC:UPI0000060779; EMBL:AL137561
A:Experimental source: adult testis; clone DKFZp434N2420
C:Genetic:
A>Note: DKFZp434N2420.1

Query Match
Best Local Similarity 23.0%; Pred. No. 1.3;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;

QY 8 HRSRCATPRGDFC-----GTERAIDQA-----SFTSMEMDTQVKGSSPL----- 50
DB 76 YRQQAOPPH---CPAEEGBGAPQALGDAPSTSVLTTAVQDVVCPLOSHALCTCCFQ 132
QY 51 ---GPAGLGAEBPAAGPQLPQSWLQPERCAVFOCAQCHAVLADSVHLAMDLSR----- 99
DB 133 PMFDRABRABQDDPRVAPQ-----QCAVC---LQPFCHLYMGCTRTGCGCLA 176
QY 100 ----SLGAVVFSRVTTNNVLLEAPFLVGEISLKGSTY-NLLFCGSCGIPVGFHLYSTHA 153
DB 177 PCCELINIGDKCLDGLVNNNSYESDILKNYLAT-RGLTWKKNL-----TES 220
QY 154 ALAALRGHFCIS-----SDKVVCYLLTKTKAIVNASSEMDIQVPLSE 194
DB 221 LVALQKGVFLSLDVRVTVGTVLCYCCGLRSRRELITVQRNRIIPASE 266

RESULT 4
A43623
-kanamycin kinase (EC 2.7.1.95) - Campylobacter jejuni
C:Species: Campylobacter jejuni
C>Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 31-Dec-2004
C:Accession: A43623
R:Tenover, F.C.; Gilbert, T.; O'Hara, P.

Plasmid 22, 52-58, 1989
A:Title: Nucleotide sequence of a novel kanamycin resistance gene, aphA-7, from Campylob

A:Reference number: A43623; MUID:89387451; PMID:2550983
A:Accession: A43623
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-250 <TEN>
A:Cross-references: UNIPROT:P4508; UNIPARC:UPI000012DEBC; GB:M29953; GB:J03316; NID:914
C:Superfamily: aminoglycoside 3'-phosphotransferase (kanamycin kinase)
C:Keywords: phosphotransferase

Query Match
Best Local Similarity 28.7%; Pred. No. 1.3;
Matches 33; Conservative 18; Mismatches 46; Indels 18; Gaps 5;

QY 105 VPSRVNNNVLEAPFLVGEISLKGSTYNLLFCGSCGIPVGFHLYSTHAALALRG-HFC 163
DB 40 IFSKTTYSVKREKEMMMWLSDKLKPDV-----IEYGREHSEYIMLSRLKHID 90
QY 164 LSSDKVVCYLLTKAIVNA-----SEMDIQVPLSEKI-AELKEKIVLTHNRKSL 213
DB 91 CFIIDHIXI---ECVNLHQLAIDINCPSSKIDVRLKELKYLIDNRADI 142

RESULT 5
A97229
NAD-dependent DNA ligase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004
C:Accession: A97229
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, J.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: A97229
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-669 <KUR>
A:Cross-references: UNIPROT:Q97FQ5; UNIPARC:UPI00000CA5D7; GB:AE001437; PIDN:AAK80620.1,
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetic:
A:Gene: CAC2673
C:Superfamily: DNA ligase (NAD), Liga type

Query Match
Best Local Similarity 27.4%; Pred. No. 6.2;
Matches 37; Conservative 22; Mismatches 45; Indels 31; Gaps 8;

QY 96 DLNR---SLGAVVFSRVTTNNVLLEAPFLVGEISLKGSTYNL-----LFCGSCGIPVGF 146
DB 366 DIRKKVYKIGSRFVFRNSDVI---PEIMGVTEETGETNEIEAPTICPGSGEIVKEGV 422
QY 147 HLSTHAALALRGHFCISLSDKRVVCYLLTKAIVNASSEMDIQVPL-LSKELAE-LKEKV 204
DB 423 HL-----FC--LNTLSCKPQVWSIVHAPASRAMNIEGFSKELTAQLPEK-- 465
QY 205 LTHNRKSLMKIUSE 219
DB 466 LNTKISIDLYRTKE 480

RESULT 6
J01977
glutamate synthase (NADH2) (EC 1.4.1.14) precursor [validated] - alfalfa
C:Species: Medicago sativa (alfalfa)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
C:Accession: J01977; PQ0551
R:Gregerson, R.G.; Miller, S.S.; Twary, S.N.; Gant, J.S.; Vance, C.P.
Plant Cell 5, 215-226, 1993
A:Title: Molecular characterization of NADH-dependent glutamate synthase from alfalfa n
A:Reference number: J01977; MUID:93200806; PMID:8453303
A:Accession: J01977

A/Molecule type: mRNA
 A:Residues: 1-2194 <GRS>
 A:Cross-references: UNIPROT:Q03460; UNIPARC:UP1000012B7FA; GB:L01660; NID:g166411; PIDN:
 A:Accession: P00551
 A/Molecule type: protein
 A:Residues: 102-114 <GRS>
 A:Cross-references: UNIPARC:UP10000172061
 C/Comment: This enzyme catalyses the reductive transfer of the amino group of glutamine
 C/superfamily: glutamate synthase (NADH/NADPH), eukaryotic type
 C/keyWords: 3fe-4s; chloroplast; iron-sulfur protein; metalloprotein; NAD; oxidoreductas
 F/1-10/Domain: propeptide #status predicted <PRO>
 F/102-2194/Product: glutamate synthase (NADH) #status experimental <MAT>
 F/102/Active site: Cys #status predicted
 F/1246,1252,1257/Binding site: 3fe-4s cluster (Cys) (covalent) #status predicted

Query Match 7.1%; Score 85; DB 1; Length 2194;
 Best Local Similarity 24.7%; Pred. No. 39;
 Matches 58; Conservative 22; Mismatches 79; Indels 76; Gaps 12;

QY 4 QPRLHRSRCATPPRGDFCGGTERAIDQAS-----FTTSMEMDTQVVKSSPLGPA 53
 DB 1037 EPLADGSR--NPKRS-----AIKQVAGSGRGVSSYYLTNADLQIKKAGAKP--- 1082
 QY 54 GLGAEPP-----AAGPOLPSWLOPERCAVFCQACQCHAVLADSVHLAMDLSR 99
 DB 1083 GEGGELPRGHKVIDIATRNSTGVLIS--PPPHDVIS-----IEDLAQIHDILKN 1133
 QY 100 SLGAVFSRVTVNNVLEAPLVGIEGSLKSTYNLLFCG-----SCGIPVG 145
 DB 1134 ANPA--ARISVLTSEAGVGIASGVKGAHVLISGHIDGSGTASRWTKIRKSAGLPWE 1190
 QY 146 FHLVSTHAALAA--LQGHFCLSSDKMVCYLKT-----KATVASEMDIQNVPL 192
 DB 1191 LGLAETHQTLVANDLGRITTLQTDGQ---LKTGRDVAIALAGAEYGFSTAPL 1241

RESULT 7
 AE2074
 Ferrichrome binding protein of ABC transporter all12147 [imported] - Nostoc sp. (strain F
 C/species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C/Accession: AE2074
 R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iritiguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A/Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium Ana
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AE2074
 A/Status: preliminary
 A/Molecule type: DNA
 A:Residues: 1-361 <KUR>
 A:Cross-references: UNIPROT:Q8V34; UNIPARC:UP100000C32E; GB:BA000019; PIDN:BA073846.1;
 A:Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: all12147
 C/Superfamily: ferrichrome-iron transport protein fecB

Query Match 7.1%; Score 84.5; DB 2; Length 361;
 Best Local Similarity 26.1%; Pred. No. 4.3;
 Matches 36; Conservative 17; Mismatches 36; Indels 49; Gaps 7;

QY 97 LSRSLGAVVF---SRVT---NNVLEAPLVGIEGSLKSTYNLLFCGSC-----GIPV 144
 DB 81 ISHALGKVKIPKPRQVRVLEENIILDSVALGVK-----PVGVVYCCQCCENFGIP 133
 QY 145 GRLVYSTHAALALRGHFLCSSDKMVCYLKTATVNASMDIQNVPLSKIALEKIV 204
 DB 134 -----SD-----LADVAVVG---NIGNOPSLKTIUSLPDL 163
 QY 205 LTHNRLSKSLMKILSEVTP 222
 DB 164 LGLTWLKKSYKILSIAP 181

RESULT 8
 D97047
 DNA ligase (NAD dependent), liga [imported] - Clostridium acetobutylicum
 C/species: Clostridium acetobutylicum
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004
 C/Accession: D97047
 R/Molling, U.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: D97047
 A/Status: preliminary
 A/Molecule type: DNA
 A:Residues: 1-663 <KUR>
 A:Cross-references: UNIPROT:Q97J88; UNIPARC:UP100000CA10F; GB:AE001437; PIDN:AAK79167.1
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C/Genetics:
 A/Gene: CAC1195
 C/Superfamily: DNA ligase (NAD), liga type

Query Match 6.9%; Score 83; DB 2; Length 663;
 Best Local Similarity 28.6%; Pred. No. 13;
 Matches 36; Conservative 18; Mismatches 44; Indels 28; Gaps 7;

QY 102 GAVVFSRVTVNNVLEAPLVGIEGSLKSTYNL-----LFCGSCGIPVGFHLYSTHAAL 155
 DB 368 GARVFLRRSNDVI--PEIMGVTEETGETKEIEAPTICPGSGEIVKEGVHL----- 417
 QY 156 MLRGHFLCSSDKMVCYLKTATVNASMDIQNV-LSKTLAE-LKEKIVLTHNRLSKSL 213
 DB 418 -----FC--EWTLCKPQMVKSIVHFAFRKAMNIEGFSEKTAQLPFR--LNIKSIIDL 467

QY 214 MKILSE 219
 DB 468 YRTKE 473

RESULT 9
 B69099
 formate hydrogenlyase, iron-sulfur subunit 2 - Methanobacterium thermoautotrophicum (st.
 C/species: Methanobacterium thermoautotrophicum
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C/Accession: B69099
 R/Smith, D.R.; Doucette-Stamm, L.A.; DeJonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N
 J. Bacteriol. 179, 7135-7155, 1997
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funci
 A/Reference number: A69000; MUID:98037514; PMID:9371463
 A/Accession: B69099
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A:Residues: 1-143 <MTH>
 A:Cross-references: UNIPROT:Q2769; UNIPARC:UP10000066781; GB:AE000929; GB:AE000666; NII
 A:Experimental source: strain Delta H
 C/Genetics:
 A/Gene: MTH1736
 C/Superfamily: nrfc protein; ferredoxin 2[4Fe-4S] homology
 F/62-118/Domain: ferredoxin 2[4Fe-4S] homology <FER2>

Query Match 6.9%; Score 82.5; DB 2; Length 143;
 Best Local Similarity 23.0%; Pred. No. 2.1;
 Matches 35; Conservative 26; Mismatches 66; Indels 25; Gaps 6;

QY 72 PERCAVFCQACQCHAVLADSVHLAMDLSRSLGAVVF-----SRVTNNVLEAPLVGIG 125
 DB 8 PELCD--ECMKERICPKNAIRVID-----GVVFCMHCSPPRAPCLNCPDAIVEVVG 60
 QY 126 SLKSTYNLLFCGSC--GIPVGFHLYSTHAALALRGHFLCSSDKMVCYLKTATVNAS 183

Db 61 AVVLEDRICGICGLCRDACPVG--AITLNERGVAVKCDLICDRKPLCVWVCPKALSES 118
 QY 184 EMDIQNVPLSEKIAELKEKIVLTHNLKSLMK 215
 Db 119 SEDM-----MAKRDKIAGELKRLKNIIMK 142

RESULT 10
 A85433
 Sugar transporter like protein [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C/Accession: A85433
 R/Anonymous: The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A/Reference number: A85001, MUID:20083488; PMID:10617198
 A/Accession: A85433
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-493 <STO>
 A/Cross-references: UNIPROT:Q23213, UNIPARC:UPI000009F22F, GB:NC_001268; NID:G1270615; F
 C/Genetics:
 A/Gene: AT4G36670
 A/Map position: 4
 C/Superfamily: glucose transport protein

Query Match 6.9%; Score 82.5; DB 2; Length 493;
 Best Local Similarity 24.9%; Pred. No. 10;
 Matches 53; Conservative 32; Mismatches 59; Indels 69; Gaps 13;

QY 79 QCAQCHAVLADSVHLMDSRSGAVVF---SRVTNNVLEAPFLVGI-----EGSLK- 128
 Db 19 QCAIVASIV--SIIFGDTGVMSGAMFIEDLKTNDVOIEV--LTGILMLCALVGSILA 74
 QY 129 -----GSTYNLFCGSC--GIPVGFHL-----YSTH 152
 Db 75 GRTSDIIGRRYITVLASILFMLGSLMGWGPVNVLLSGCTGLGVFLMVAPVYSAE 134
 QY 153 AALALALRG-----HFLSSDKAVCYLTK---TKAIVNAS---EMDIQNV---LSKIA 197
 Db 135 IATASHHGLASLPHLCISIGILGYIVNYFFSLPHMIGWRMLGLIAVPSLVLAEGIL 194
 QY 198 ELKE--KIVLTHNLKSLMKILSEV--TPDSK 226
 Db 195 KMPESPRLIMQGRKKEKILVNSPEBAE 227

RESULT 11
 F87678
 DNA polymerase I [imported] - Caulobacter crescentus
 C/Species: Caulobacter crescentus
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C/Accession: F87678
 R/Nierman, W.C.; Deibyl, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; Debby, R.T.; Dodson, R.J.; Durrin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of Caulobacter crescentus.
 A/Reference number: A87249; MUID:21173698; PMID:11259647
 A/Accession: F87678
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-967 <STO>
 A/Cross-references: UNIPROT:Q9A2U2, UNIPARC:UPI00000CTAB1, GB:AE005673; NID:G13425184; F
 C/Genetics:
 A/Gene: CC3464
 C/Superfamily: DNA-directed DNA polymerase I

Query Match 6.8%; Score 82; DB 2; Length 967;
 Best Local Similarity 24.1%; Pred. No. 26;
 Matches 51; Conservative 27; Mismatches 72; Indels 62; Gaps 11;

QY 12 CATPRGDFCGGTERAIDQ---ASFTTSMENDT---QVVGSSPLGAPAGAEBAAGQ 65
 Db 271 CDTPLPQPLDALTLVPEPDKEALAAFLQWEPFSLARVGDGSAATPGTL--DRPAAPK 328
 QY 66 LP---SWL-----OPERC--AVFQCAQCHAVLADSVHLMDSRSGAVVF 108
 Db 329 APVSVSVYMGAAARAAHPVEPEVKIDHAAVACVRDLATLKAVVAKATD---KGLAFT 384
 QY 109 VTNNVVLEAPFLVIGESLKGSTYNLFCGSCGIPVGFHLVSTHAALALRGHCLSDPK 168
 Db 385 ETD-----ALSSAT-----AGLCGV-----SLAIPBEACTIP-- 412

QY 169 MCVYLKTKAIVNASEMDIQNVPLSEKIAELK 200
 Db 413 -ISHCKADGLAFEAPADIEQIPLDVITLTK 443

RESULT 12
 D90404
 Transport protein, probable [imported] - Sulfolobus solfataricus
 C/Species: Sulfolobus solfataricus
 C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
 C/Accession: D90404
 R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
 arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to Genbank, April 2001
 A/Description: Sulfolobus solfataricus complete genome.
 A/Reference number: A99139
 A/Accession: D90404
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-329 <RUR>
 A/Cross-references: UNIPROT:Q97W97, UNIPARC:UPI00000646FA; GB:AE006641; NID:G13815639; F
 C/Genetics:
 A/Gene: SS02338

Query Match 6.8%; Score 81.5; DB 2; Length 329;
 Best Local Similarity 22.2%; Pred. No. 7.4;
 Matches 40; Conservative 30; Mismatches 47; Indels 63; Gaps 7;

QY 87 LADSVHLA-WD-----LSRSLGAVVPSRYTNNV-----LE 116
 Db 33 LSSSHMLAVWEPALVLPFLGRIIGSFYOVFKNSVISCPEPLGFLVILONFLGALL 92
 QY 117 APFLVG-----IEGSLKSTYNLFCGSCGIPVGFHLVSTHAALALRGHCLAS 166
 Db 93 VRFVLGVIIFGLTSYAVESAVKSGRNVLVGFTTAGVPIGVN1----- 134

QY 167 DKAVCY-LIKTKAIVNASEMDIQNVPLSE---KIAELKEKIVLTHNLKSLMKILSEVP 222
 Db 135 -SYAVVLLKNNMNVINISGILIMLALFELNGEPFERSKISVSPRLSILYVALRP 193

RESULT 13
 H85767
 glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C/Species: Escherichia coli
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C/Accession: H85767
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe-
 iller, L.; Grobeck, E.J.; Davis, N.W.; Llim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A/Reference number: A85480; MUID:21074935; PMID:11206551
 A/Accession: H85767
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-457 <STO>
 A/Cross-references: UNIPROT:Q8X673; UNIPARC:UPI00000DOB5; GB:AE005174; NID:G12515601; F
 A/Experimental source: strain O157:H7, substrain EDL933
 C/Genetics:
 A/Gene: uidB

C:Superfamily: melibiose carrier protein

```

Query Match          6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 11;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSLSGAVVFSRVNTNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 WSLPVALVALALASIGGVMTVMMALEADTVEXEYGLTGVR--IEGLTYSLSFSTRKCG 374
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 142 IPVGFHLVSTHALLALRGHFC--LSSDKVVCYLTKKALVNASMDIQNV-----PLSE 194
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 375 QAIG---GSIPAFILGSGYIANOVQTPPEVIMGIRTSIALVPCGFMLAFVIIMFYPLTD 431
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 195 KIAELKEKIVLTNNRLKSLMKILSEVT 221
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 432 K--KFKELIVEIDNRKKVQOQLISDIT 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14

glucuronide permease [imported] - Escherichia coli (strain O157:H7, substrain R1MD 05095 C90919)
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: C90919
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gaeawara, N.; Yasunaga, T.; Kubara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: C90919
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-457 <HAY>
 A:Cross-references: UNIPROT:Q8X673; UNIPARC:UPI000000DBE5; GB:BA000007; P1DN:BA035746.1
 A:Experimental source: strain O157:H7, substrain R1MD 050952
 C:Genetics:
 A:Gene: ECG2323
 C:Superfamily: melibiose carrier protein

```

Query Match          6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 11;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSLSGAVVFSRVNTNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 WSLPVALVALALASIGGVMTVMMALEADTVEXEYGLTGVR--IEGLTYSLSFSTRKCG 374
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 142 IPVGFHLVSTHALLALRGHFC--LSSDKVVCYLTKKALVNASMDIQNV-----PLSE 194
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 375 QAIG---GSIPAFILGSGYIANOVQTPPEVIMGIRTSIALVPCGFMLAFVIIMFYPLTD 431
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 195 KIAELKEKIVLTNNRLKSLMKILSEVT 221
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 432 K--KFKELIVEIDNRKKVQOQLISDIT 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 15

glucuronide permease uidB - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C:Accession: B64918
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: B64918
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-457 <BLAT>
 A:Cross-references: UNIPROT:P30868; UNIPARC:UPI0000137B23; GB:AE000257; GB:U00096; NID:9

A:Experimental source: strain K-12, substrain MGL655

C:Genetics:

A:Gene: uidB; gusB

C:Superfamily: melibiose carrier protein

C:Keywords: carrier protein; transmembrane protein

F:37-53/Domain: transmembrane #status predicted <TM01>

F:81-97/Domain: transmembrane #status predicted <TM02>

F:152-168/Domain: transmembrane #status predicted <TM03>

F:184-200/Domain: transmembrane #status predicted <TM04>

F:231-247/Domain: transmembrane #status predicted <TM05>

F:263-279/Domain: transmembrane #status predicted <TM06>

F:310-326/Domain: transmembrane #status predicted <TM07>

F:408-424/Domain: transmembrane #status predicted <TM08>

```

Query Match          6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 11;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;

QY 95 WDLRSLSGAVVFSRVNTNV-----VLEA-----PFLVGIEGSLKSTYNLL-FCGSCG 141
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 WSLPVALVALALASIGGVMTVMMALEADTVEXEYGLTGVR--IEGLTYSLSFSTRKCG 374
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 142 IPVGFHLVSTHALLALRGHFC--LSSDKVVCYLTKKALVNASMDIQNV-----PLSE 194
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 375 QAIG---GSIPAFILGSGYIANOVQTPPEVIMGIRTSIALVPCGFMLAFVIIMFYPLTD 431
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 195 KIAELKEKIVLTNNRLKSLMKILSEVT 221
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 432 K--KFKELIVEIDNRKKVQOQLISDIT 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: May 18, 2006, 15:41:35
 Job time : 40 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2006, 15:33:31 ; Search time 299 Seconds

(without alignments)
708.457 Million cell updates/sec

Title: US-09-942-052a-728

Perfect score: 1198

Sequence: 1 MAOPLRHRSRCATPPRGDF.....LKSIMKILSEVTPDQSKPEN 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1198	100.0	229	1 OIP5 HUMAN	O43482 homo sapien
2	122	10.2	218	2 O9CKR6 MOUSE	O9CKR6 mus musculu
3	118	9.8	462	2 O4PAB5 USMMA	O4PAB5 usitlago ma
4	115.5	9.6	204	2 O9CZJ6 MOUSE	O9CZJ6 m 11 days e
5	113	9.4	233	1 CU045 HUMAN	O9NYP9 homo sapien
6	113	9.4	233	1 O54220 HUMAN	O54220 homo sapien
7	110.5	9.2	232	1 CU045 PANTR	O68UR5 pan troglod
8	109	9.1	207	2 O5BLBI BRABE	O5BLBI brachydanio
9	107	8.9	532	2 O5AH16 DICDI	O5AH16 dicystoselei
10	105	8.8	155	1 YCZC SCHPO	O9B802 schizosacch
11	96	8.0	391	2 O33B05 ORYSA	O33B05 oryza sativ
12	96	8.0	497	2 O3V0J4 MOUSE	O3V0J4 mus musculu
13	94	7.8	1556	2 O4O626 LEIMA	O4O626 leishmania
14	94	7.8	702	2 O2OSG3 ORYSA	O2OSG3 oryza sativ
15	92	7.7	361	2 O8S611 ORYSA	O8S611 oryza sativ
16	92	7.7	695	2 O9CH87 LACIA	O9CH87 lactococcus
17	91.5	7.6	410	2 O2JZY7 RHET	O2JZY7 rhizobium e
18	91	7.6	628	2 O69KE0 ORYSA	O69KE0 oryza sativ
19	90.5	7.6	272	2 O4OQJ4 DESAC	O4OQJ4 desulfurimo
20	90.5	7.6	597	2 O83693 TREBA	O83693 treponema p
21	89.5	7.5	502	2 O5OPD2 ENTHI	O5OPD2 entamoeba h
22	89.5	7.5	661	2 O5OV43 ENTHI	O5OV43 entamoeba h
23	89.5	7.5	661	2 O5OW28 ENTHI	O5OW28 entamoeba h
24	89.5	7.5	2034	2 O4E556 TRYCR	O4E556 trypanosoma
25	89	7.4	633	2 O5M052 STRT1	O5M052 streptococc
26	89	7.4	664	1 CHPR HUMAN	O5ME91 homo sapien
27	89	7.4	735	2 O5MAR3 STRT2	O5MAR3 streptococc
28	89	7.4	893	1 POL2_BAMMA	O6S329 barley mild
29	88.5	7.4	152	2 O5OXW7 ENTHI	O5OXW7 entamoeba h
30	88.5	7.4	384	2 O3A3W4 PELCD	O3A3W4 pelobacter
31	88.5	7.4	638	2 O75RZ6_CIOIN	O75RZ6 clona intes

32	88	7.3	250	1 KK47 CAMJE	P14508 campylobact
33	88	7.3	493	2 O68VH2 IPOTIA	O68VH2 barley mild
34	87.5	7.3	634	2 O6CX82 KLUYA	O6CX82 kluyveromyc
35	87.5	7.3	663	2 O3UZ33 MOUSE	O3UZ33 m nod-deriv
36	87.5	7.3	696	2 O4UTS4_CORK	O4UTS4 corynebacte
37	87	7.3	524	2 O3UTD8_MOUSE	O3UTD8 mus musculu
38	87	7.3	592	2 O3UGJ9_MOUSE	O3UGJ9 mus musculu
39	87	7.3	664	1 CHPR MOUSE	O3U49 mus musculu
40	87	7.3	664	2 O3U49_MOUSE	O3U49 mus musculu
41	86.5	7.2	669	2 O97RQ5_CLOAB	O97RQ5 clostridium
42	86.5	7.2	2093	2 O5KXK3_9RAB	O5KXK3 istahan vir
43	86	7.2	145	2 O839T4_ENTFA	O839T4 entercoccu
44	86	7.2	327	2 O4NE64_9MIC	O4NE64 arthrobacte
45	86	7.2	331	2 O6GXK5_COCCA	O6GXK5 cochlidiobol

ALIGNMENTS

RESULT 1
OIP5_HUMAN STANDARD; PRT; 229 AA.
ID O43482; O96BX7;
AC O43482; O96BX7; Integrated into UniProtKB/Swiss-Prot.
DT 03-APR-2002, sequence version 2.
DT 07-FEB-2006, entry version 29.
DE Opa-interacting protein 5.
GN Name=OIP5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=98125741; PubMed=9466265;
RA Williams J.M., Chen G.-C., Zhu L., Rees R.F.;
RT "Using the yeast two-hybrid system to identify human epithelial cell
RT proteins that bind gonococcal Opa proteins: intracellular gonococci
RT bind pyruvate kinase via their Opa proteins and require host pyruvate
RT for growth.";
RT Mol. Microbiol. 27:171-186(1998).
RL [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RS TISSUE=uterus;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapietron M., Soares M.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abrahams R.D., Mullany S.J.,
RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hellon E., Kettelman M., Madan A.C., Shevchenko Y., Bouford G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Rodriguez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherach A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL -1- SUBUNIT: Binds outer membrane protein Opa from Neisseria
RL gonorrhoeae.
CC -1- INTERACTION:
CC P04049:RAFI; NbExp=3; IntAct=EBI-536879, EBI-365596;
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license

```
CC -----
DR EMBL: AF025441, AAC39561.1, ALT_INT, mRNA.
DR EMBL: BC015050, AAH15050.1, -, mRNA.
DR IntAct: O43482, -.
DR Ensembl: ENSG00000104147, Homo sapiens.
DR H-invdB: HIX0012150, -.
DR HGNC: HGNC:20300, OIP5.
DR MIM: 606020, gene.
DR LinkHub: O43482, -.
DR GO: GO:0005515, F:protein binding; TAS.
DR GO: GO:0007154, P:cell communication; NAS.
FT CHAIN 1 229 /Frid=PRO_0000058038.
SQ SEQUENCE 229 AA; 24691 MW; 0EBD4006193A3106 CRC4;

Query Match 100.0%; Score 1198; DB 1; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.3e-104;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPQLRHRSRCATPPRGDPFCGGTBEAIDQASPTTSMENDTQVVGKSSPLGPAGLGAEEP 60
DB 1 MAAPQLRHRSRCATPPRGDPFCGGTBEAIDQASPTTSMENDTQVVGKSSPLGPAGLGAEEP 60
QY 61 AAGPOLPSWLQPERCAVFCQACQAVLADSVHLAMDLSRSLGAVFSRVYNNVLEAPPL 120
DB 61 AAGPOLPSWLQPERCAVFCQACQAVLADSVHLAMDLSRSLGAVFSRVYNNVLEAPPL 120
QY 121 VGIEGSLKSGSTNNLFCGSGGIVGPHLYSTHAALALRHFCLSSDQKVCYLTKKATV 180
DB 121 VGIEGSLKSGSTNNLFCGSGGIVGPHLYSTHAALALRHFCLSSDQKVCYLTKKATV 180
QY 181 NASEMDIONVPLESEKTAELKEKIVLTHNRLKSLMKLTISEVTPQSKPEN 229
DB 181 NASEMDIONVPLESEKTAELKEKIVLTHNRLKSLMKLTISEVTPQSKPEN 229

RESULT 2
O9CKR6_MOUSE PRELIMINARY; PRT; 218 AA.
AC O9CKR6,
DT 01-JUN-2001, Integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE 13 days embryo head cDNA, RIKEN full-length enriched library,
clone:3110025H3 product:hypothetical protein, full insert sequence.
GN Name=261003C1ORik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Head; STRAIN=C57BL/6J;
RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Fritsch M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilmink L.G., Aldins V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Auraliyya R.N., Bailey T.L.,
RA Banasi M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutierbuck D.R.,
RA Crowe M.U., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Guestinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapi A., Katoh M., Kawasawa Y., Kello J., Kitamura H.,
RA Kitano H., Kollipati G., Krieten S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Larau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Moutagu-Tabar S., Mulder N., Nakano N., Nakaguchi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pebole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Roost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yang K.,
RA Yamashita H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Iannura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Kono H., Nakano K., Nimmiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Wataniki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA PubMed=16141073; DOI=10.1126/science.1112009;
RA RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi U., Bono H., Kondo S.,
RA Nishio T., Ose T., Saito R., Suzuki H., Yamashita H., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schram L.M., Kanapi L.M., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Brusic V., Chothia K., Corbani L.E., Coullins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guestinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Malais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pereira G., Pebole G.,
RA Petrovsky N., Pillai R., Pontius U.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynnaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirokawa-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Haseizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
```

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batciov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Peeble G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection.";
RA Nature 409:685-690(2001).
[6]
RA NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J; TISSUE=Head;
RA MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RA "Normalization and subtraction of cap-trapper-selected cDNAs to
RA prepare full-length cDNA libraries for rapid discovery of new genes.";
RA Genome Res. 10:1517-1530(2000).
[7]
RA NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J; TISSUE=Head;
RA MEDLINE=2050913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Ozawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RA "RIKEN integrated sequence analysis (RISA) system-34-format
RA sequencing pipeline with 384 multicapillary sequencer.";
RA Genome Res. 10:1757-1771(2000).
[8]
RA NUCLEOTIDE SEQUENCE.
RA STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Koude M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
CC EMBL: AK014084; BAB29147.1; -; mRNA
CC
CC Ensembl: ENSMUSG00000022878; Mus musculus.
CC
CC MGI: MGI:1913828; 2610039C10R1X.
CC
CC Hypothetical protein.
CC
CC SEQUENCE 218 AA; 24447 MW; 553905BFDDE8C2 CRC64;
QY
QY Query Match 10.2%; Score 122; DB 2; Length 218;
QY Best Local Similarity 23.3%; Pred. NO. 0.0065;
QY Matches 48; Conservative 30; Mismatches 102; Indels 26; Gaps 6;
QY 29 DQAFPTSMWDTQVVGSSPLGAGIAGEPAAGPOLPSWLOPERCAVFOCAQCHAVLA 88

DB 15 DSSRYLRLQWMA--MSSADALGLEKEREPEKKAANP-----LVFLCARCRPLG 64
QY 89 DSHLAMDLSR-SLGAVERRVNNVYLAEPVLVIGLSLKGSTYNNLPGSGGIPVGFH 147
DB 65 DS--LTVWVSAQEDNTCLILRSVCNVSVDKPKLSKCRDBDGCILEALYCTGSLSGY 122
QY 148 LYSTHALALRGHFCISDPKWCYLL-----KTKALVNAEMDQNPSEKIAE 198
DB 123 YRCTPKRLDVKRDLFCISVTVESYTLGSGSEKQIVSKDKLFLNE----SRVELKSIKQ 178
QY 199 LKEKIVLTNRKLSLKNKILSEVTPDQ 224
DB 179 MEVLTLQKKLKEVESKLSLAQPGQ 204
RESULT 3
ID Q4PAB5 USTWA PRELIMINARY; PRT; 462 AA.
AC Q4PAB5;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=UM02948.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=521;
RA Birren B.W., Nusbaum C., Abebe A., Abouelleil A., Adekoya B.,
RA Ait-Zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H.M., Armbruster J., Bachanteng P., Baldwin J., Barry A.,
RA Bayul T., Blithstein B., Bloom T., Bye J., Boguslavsky L.,
RA Borowsky M., Boukhalter B., Brunahe A., Butler J., Calixte N.,
RA Calvo S.E., Camarata J., Campo K., Chang J., Cheahsang Y.,
RA Cifredo M., Collamore A., Considine T., Cook A., Cooke P., Corum B.,
RA Clomo C., David R., Dawoe T., Degray S., Dodge S., Dooley K.,
RA Dorje P., Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T.,
RA Engels R., Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgeraid M., Foley K., Gage D., Galagan J.E., Geartin G., Gnerre S.,
RA Galtke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysellis M., Karlsson E.,
RA Kellis C., Kieu A., Klesner P., Kodira C., Kulbokas E., Labutic K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-Toh K., Liu X., Lokysiang T., Lokysiang Y., Lucien O.,
RA Lui A., Ma L.-O., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Maucelli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrum J., Menes L.,
RA Mesitov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C.B., Nizzari M., Norbu C.,
RA Norbu N., O'Donnell P., Okawo O., O'Leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramassamy R., Rameau R., Ray V., Raymond C.,
RA Ratta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settillai S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,
RA Spencer B., Stalker J., Stange-Thomann N., Stavropoulos S.,
RA Steeton K., Stone C., Stone S., Stubbs M., Talamas J., Thuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thouloutang Y., Topham K.,
RA Towey S., Tsamila T., Tsono N., Vallée D., Vaesliliev H.,
RA Venkataraman V.S., Vishon J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wu Y., Yaman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
RA Zimmer A., Zody M., Zander E.S.;
RT "The genome sequence of Ustilago maydis."
RT Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 DR EMBL; AACPO1000101; EAK84120.1; -; Genomic_DNA.
 KW Hypothetical protein
 SQ SEQUENCE 462 AA; 49283 MW; EDDCPDF1757B509 CRC64;
 Query Match 9.8%; Score 118; DB 2; Length 462;
 Best Local Similarity 25.9%; Pred. No. 0.04;
 Matches 64; Conservative 28; Mismatches 109; Indels 46; Gaps 10;
 QY 2 AAGPLAHRRCATPPGCGCGTGERALDQASFTTSNEMDQV-VKSSPLGPGAGLGAEP 60
 DB 109 AAGPLSSDDDDMPPKQ-----GSLTFTT---STRARGRGGRGPRGRGRG 153
 QY 61 AA-----GPGV-PSWLPERCAYPQCAQCAVLADSV-HLAMPDSRLGAVFESRVT 110
 DB 154 AGSTTVVHKPDLBESDSEEPNPLVFQCCCRLLGLDSLAFVATDI--DLGIVILSDVS 211
 QY 111 NNVLBAPPLVGLIE-GSLKGSYVNLFFCGSGCPIPVGFHLVSTHAALALRGHECLSSDKM 169
 DB 212 EIIQODTVEYESTEPGKDIGSTFARLRACACNAVAGRNRYTTPRDDLDLDFCSLEVDAL 271
 QY 170 VCVLLTKTAIVNASMDIO-----NPLSEKLAELKEKIVLH-NRLKS 212
 DB 272 YTYQLGSNNTYRQKEDEDDQDALVDGAGSKPARASRPHETHTVQAEEDTKVLLTKMERTRA 331
 QY 213 LMKILSE 219
 DB 332 LTFELSD 338
 RESULT 4
 Q9CZJ6_MOUSE PRELIMINARY; PRT; 204 AA.
 ID Q9CZJ6_MOUSE
 AC Q9CZJ6;
 DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2001, sequence version 1.
 DT 07-FEB-2006, entry version 21.
 DE 11 days embryo whole body cDNA, RIKEN full-length enriched library,
 DE clone:2700078124 product:RIKEN cDNA 2610039C10 (2 cells egg cDNA,
 DE RIKEN full-length enriched library, clone:B02001B09 (2610039C10R1K
 DE product:hypothetical protein, full insert sequence)
 DE protein).
 DE Name=2610039C10R1K;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Egg, and Whole body;
 RC PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilmot L.G., Altshuler S., Allen J.E.,
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Banerji M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,

RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Geislich-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummel L., Iacono M., Ikeno K., Iwama A., Ishikawa T.,
 RA Jakt A., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura T.,
 RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.P., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Moutareau-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.U., Pavoni G., Peesle G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugura K., Sultana R., Takenawa Y., Taki K.,
 RA Tammaja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yang K.,
 RA Yamashita H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Martick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukushima S., Kanamori-Katayama M., Suzuki M., Aoki U., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimmiya N.,
 RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Wataniki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome";
 RL Science 309:1559-1563(2005).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Egg, and Whole body;
 RC PubMed=16141073; DOI=10.1126/science.1112009;
 RG RIKEN Genome Exploration Research Group, and Genome Science Group
 RG (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense transcription in the mammalian transcriptome";
 RL Science 309:1564-1566(2005).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CS7BL/6J; TISSUE=Egg, and Whole body;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1098/nature01266;
 RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakai K., Otsu N., Saito R., Suzuki H., Yamashita H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Balderelli R., Hill D.P., Bull C., Hume D.A., Quackenbush J.,
 RA Schirini L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brusic V., Chotina C., Corbett L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierzki R.M., King B.L.,
 RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltale L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.U., Perle G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenawa Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Satake N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Iishi Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yashunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Egg, and whole body;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batelov S., Casavant T.,
 RA Fleischman W., Gaasterland T., Gissi C., King B., Kocha H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Perole G., Quackenbush J.,
 RA Schirrl L.M., Stadl P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Futuro M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaets P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690(2001).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Egg, and whole body;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RN Genome Res. 10:1617-1630(2000).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Egg, and whole body;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasahagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanaka M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RN Genome Res. 10:1757-1771(2000).
 RN [8]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Futuro M.,
 RA Hangaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurikawa C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito R., Saito R., Sakai C., Sakai T.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa Y., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RN Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Egg;
 RX Arakawa T., Carninci P., Fukuda S., Haehizume W., Hayashida K.,
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
 RA Kawai J., Kojima M., Kono H., Murata M., Nakamura M., Ninomiya N.,
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watanaka A.,
 RA Muramatsu M., Hayashizaki Y.;
 RN Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6; TISSUE=Eye;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Krausberg R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Urdin T.B., Tohyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarnate P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 Query Match 9.6%; Score 115.5; DB 2; Length 204;
 Best Local Similarity 25.6%; Pred. No. 0.024;
 Matches 50; Conservative 29; Mismatches 97; Indels 19; Gaps 7;
 QY 29 DQAFPTTSMEDTQVYKSSPLGAGLGAEPAPGQPSWLOPERCAVFOCAQCHAVLA 88
 Db 15 DSSRYRLQKWMAN--WSSADALGLEKREPEREKAAANP-----LVFLCARCRPLG 64
 QY 89 DSVHLAMWLSR-SLGAVFSSRVNNVYVLEAPFLVIGSLKSGTYNLLFCGSGCIPYGFH 147
 Db 65 DS--LTWVAISOEDTNCILLRVSVCNVSVDPEPLSKCRDBDCGILEALYCTGCSLSIGVY 122
 QY 148 LVSTHAALALRGHPCISSDQXVCYLL--KTRIVNASEMDIONVPLSEIATLKEIVL 205
 Db 123 YRCTPPKLDYKRLPCLISVAEVSYYLGSSEKQIV-SEDKLPNL--BSRVEIKSIKQ 178
 QY 206 THNRLKSLMKILSEV 220
 Db 179 MEEVLTALQKUREV 193
 RESULT 5
 CU045 HUMAN STANDARD; PRT; 233 AA.
 ID CU045 HUMAN
 AC G9NVP5;
 DT 11-JUN-2001, integrated into UniProtKB/Swiss-Prot.
 DT 01-OCT-2000, sequence version 1.
 DT 07-FEB-2006, entry version 27.
 DE Protein C21orf45.
 GN Name=C21orf45;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
 OC Homo.
 NCBI_TaxID=9606;
 [1]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RX MEDLINE=20237674; PubMed=10773462; DOI=10.1016/S0378-1119(00)00089-5;
 RA Slavov D., Hattori M., Sakaki Y., Rosenthal A., Shmizu N.,
 RA Minoshima S., Kudoh J., Yaspo M.-L., Ramser J., Reinhardt R.,
 RA Reimer C., Clancy K., Rynditch A., Gardiner K.;
 RT "Criteria for gene identification and features of genome organization:
 RT analysis of 6.5 Mb of DNA sequence from human chromosome 21.";
 RN Gene 247:215-232(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Blood;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Krausberg R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Urdin T.B., Tohyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarnate P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

```

RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywicki M.T., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.U.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SIMILARITY: To S.pombe SPC970.12.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AF231921; AAF72945.1; -; mRNA.
DR EMBL: BC042917; AAH42917.1; -; mRNA.
DR Ensembl: ENSG00000159055; Homo sapiens.
DR HENC: HENC:1286; C21orf45.
FT CHAIN 1 233 Protein C21orf45.
FT /FTID=PRO_0000079515.
SQ SEQUENCE 233 AA; 25863 MW; A07522806C4B6221 CRC64;

Query Match 9.4%; Score 113; DB 1; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.05;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTTSMENDTVQVKGSSPLGPAAGLGAEPAPGQPLSMLOPRCAVFOCAQCHAVLADS 90
DB 49 ASWMSMSSEASV---ADMERAOQL--EEBAALAE-----ERPLVFLSCGCRRLPGDS 95
QY 91 VHLAMPDLR-SLGAVFVSRTNNVLEAPFLVIGESLGKSTYNLLFCSCGCIPIVGFHLX 149
DB 96 --LSWVASQEDTNCILRCVSCNVSVDKQKLSKREKNGCVLETICACGCSINLGVYR 153
QY 150 STHAALALRGHFCSSDDKAVCYLL--KTKAIVNASMDIONVPLSEKIAELKEKIVLTH 207
DB 154 CTPKNDLYKRDPLCLSVLAESVYLGSSEKQIV--SEDKELFNL--ESRVEIKSLTQME 209
QY 208 NRKLSLMKILSE 219
DB 210 DVKALQMKLWE 221
QY 208 NRKLSLMKILSE 219
DB 210 DVKALQMKLWE 221

RESULT 6
Q54220 HUMAN PRELIMINARY; PRT; 233 AA.
ID 054220_HUMAN
AC 054220
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DE 07-MAR-2006, entry version 8.
DE Hypothetical protein FLJ90800 (FAPPI-associated protein 1).
GN Name=FAPPI;
OS Homo sapiens (human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=16303743; DOI=10.1093/dnares/12.2.117;
RA Otsuki T., Ota T., Nishikawa T., Hayashi K., Suzuki Y., Yamamoto J.,
RA Makamatsu A., Kimura K., Sakamoto K., Hatanaka N., Kawai Y., Ishii S.,
RA Saito K., Kojima S., Sugiyama T., Ota T., Okano K., Yoshikawa Y.,
RA Aotsuka S., Saeki N., Hattori A., Okumura K., Nagai K., Sugano S.,
RA Isegaki T.,
RT "Signal sequence and keyword trap in silico for selection of full-
RT length human cDNAs encoding secretion or membrane proteins from oligo-
RT capped cDNA libraries."

```

```

RL DNA Res. 12:117-126(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=21061410; PubMed=11001876; DOI=10.1042/0264-6021.3510019;
RA Dowler S.J., Currie R.A., Campbell D.G., Deak M., Kular G.,
RA Downes C.P., Alessi D.R.;
RT "Identification of pleckstrin-homology-domain-containing proteins with
RT novel phosphoinositide-binding specificities."
RL Biochem. J. 351:19-31(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Dowler S.J.;
RT "Identification of FAPPI interacting proteins."
RL Submitted (JUN-2001) to the EMBL/GenBank/DDA databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AK075281; BAC11517.1; -; mRNA.
DR EMBL: AF387845; AAK70498.1; -; mRNA.
DR Ensembl: ENSG00000159055; Homo sapiens.
SQ SEQUENCE 233 AA; 25863 MW; A07522806C4B6221 CRC64;

Query Match 9.4%; Score 113; DB 2; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.05;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY 31 ASFTTSMENDTVQVKGSSPLGPAAGLGAEPAPGQPLSMLOPRCAVFOCAQCHAVLADS 90
DB 49 ASWMSMSSEASV---ADMERAOQL--EEBAALAE-----ERPLVFLSCGCRRLPGDS 95
QY 91 VHLAMPDLR-SLGAVFVSRTNNVLEAPFLVIGESLGKSTYNLLFCSCGCIPIVGFHLX 149
DB 96 --LSWVASQEDTNCILRCVSCNVSVDKQKLSKREKNGCVLETICACGCSINLGVYR 153
QY 150 STHAALALRGHFCSSDDKAVCYLL--KTKAIVNASMDIONVPLSEKIAELKEKIVLTH 207
DB 154 CTPKNDLYKRDPLCLSVLAESVYLGSSEKQIV--SEDKELFNL--ESRVEIKSLTQME 209
QY 208 NRKLSLMKILSE 219
DB 210 DVKALQMKLWE 221
QY 208 NRKLSLMKILSE 219
DB 210 DVKALQMKLWE 221

RESULT 7
CU045_PANTR STANDARD; PRT; 232 AA.
ID CU045_PANTR
AC 0680T5;
DT 23-NOV-2004, integrated into UniProtKB/Swiss-Prot.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Protein C21orf45 homolog.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pan.
OX NCBI_TaxID=9598;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15164055; DOI=10.1038/nature02564;
RA Matanabe H., Fujiyama A., Hattori M., Taylor T.D., Toyoda A.,
RA Kuroki Y., Noguchi H., Benkhalil A., Leirach H., Sudbrak R., Kube M.,
RA Tienzer S., Galgoczy P., Platzer M., Schafke M., Nordstark G.,
RA Bioecker H., Hellmann I., Khaitovich P., Paabo S., Reinhardt R.,
RA Zheng H.-J., Zhang X.-L., Zhu G.-F., Wang B.-F., Fu G., Ren S.-X.,
RA Zhao G.-P., Chen Z., Lee Y.-S., Cheong J.-E., Choi S.-H., Wu K.-M.,
RA Liu T.-T., Heiso K.-J., Teal S.-F., Kim C.-G., Oota S., Kitano T.,
RA Kohara Y., Saitou N., Park H.-S., Wang S.-Y., Yaapo M.-L., Sakaki Y.;
RT "DNA sequence and comparative analysis of chimpanzee chromosome 22."
RT Nature 429:382-388(2004).
CC -1- SIMILARITY: To S.pombe SPC970.12.

```



```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL, AL954205; CAH18576.1; -; Genomic DNA.
FT CHAIN 1 232 Protein C21orf45 homolog.
FT FTID=PRO_0000079516.
SQ SEQUENCE 232 AA; 25832 MW; 65498BB369245E7 CRC64;

Query Match 9.2%; Score 110.5; DB 1; Length 232;
Best Local Similarity 27.1%; Pred. No. 0.086;
Matches 52; Conservative 29; Mismatches 88; Indels 23; Gaps 8;

QY 31 ASFTTSEMDTQVWKGSSPLGPAGLGAEPAPGQLPSWLOPERCAVFOCAQCHAVLADS 90
DB 49 ASWMSMSSEBASY-----ADMERARL--EEMAAA-----EERPLVFLSCGRRLPGDS 94
QY 91 VHLAMDLSR-SLGAVFVSRTNNVLEAPPLVGIKSLKSTYNLLFCGSCGIPVGFHLY 149
DB 95 --LSWVASQEDTNCILRLRCVSCNVSVDKQKLSKREKNGCVLETLCCAGCSINLGVYR 152
QY 150 STHAALALRGHFCLSKDKVICYLL--KTAIVNASMDIQNPVLSKIMLEKIVLTH 207
DB 153 CTFKRLDYKDLFCLSVEALIESYLVGSSEKQIV-SEDKELFNL---ESRVEIERKSLTQME 208
QY 208 NRLKSLMKITSE 219
DB 209 DVLKALQMKLME 220

RESULT 8
OSBLB1_BRARE PRELIMINARY; PRT; 207 AA.
AC OSBLB1_BRARE
DT 12-APR-2005, integrated into UniProtKB/TrEMBL.
DT 12-APR-2005, sequence version 1.
DE LOC553502, entry version 6.
DE LOC553502 protein (Fragment).
GN Name=LOC553502;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang U., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.V., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smilg D.E.,
RA Scherch A., Schein J.E., Jones S.J.W., Maira M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RG NIH MGC Project;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.

```

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL, BC090521; AAH90521.1; -; mRNA.
FT CHAIN 1 111 Protein C21orf45 homolog.
FT FTID=PRO_0000079516.
SQ SEQUENCE 207 AA; 22941 MW; AFI10ACCF375A28C CRC64;

Query Match 9.1%; Score 109; DB 2; Length 207;
Best Local Similarity 25.5%; Pred. No. 0.1;
Matches 37; Conservative 28; Mismatches 66; Indels 14; Gaps 6;

QY 76 AVFCAQCHAVLADSVHIAWDL-SRLGAVFVSRTNNVLE-APPLVGIKSLKSTYN 133
DB 54 AVFCAQCHAVLADSVHIAWDL-SRLGAVFVSRTNNVLE-APPLVGIKSLKSTYN 111
QY 134 LIFCGSGCIPVGFHLYSTHAALALRGHFCLSKDKVICYLLKTAIVNASMDIQNPVLS 193
DB 112 LT-CRGCSESLGVWYISTPKLDCKSLFCFVFNENIESYVSGPG-QQWELDRDQKPT 169
QY 194 -----EKIAELKEKIVLTHNRL 210
DB 170 LEYQDTVHQQMTBITSLAVLIGRL 194

RESULT 9
OS4H16_DICD1 PRELIMINARY; PRT; 532 AA.
AC OS4H16_DICD1
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DE LOC553502, entry version 3.
DE Hypothetical protein.
GN ORFNames=DD80188417;
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Eukaryota; Dictyostelida; Dictyostelium.
OC NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RX PubMed=15875012; DOI=10.1038/nature03481;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.A.,
RA Sugand R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Kontorov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pillcher K., Chen G., Saunders D., Sodergren E.J., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Baeson N.,
RA Farbrother P., Desany B., Just E., Morio T., Roat R., Churcher C.M.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D.M., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K.D., Quiles W., Madan Babu M., Salto T.,
RA Buchrieser C., Wardrop A., Felder M., Thangavelu M., Johnson D.,
RA Knights A., Loulèsed H., Mungall K.L., Oliver K., Price C.,
RA Quail M.A., Urushihara H., Hernandez J., Rabinowitch E., Steffen D.,
RA Sanders M., Ma J., Kohara Y., Sharp S., Simmonds M.N., Spiegler D.,
RA Tivey A., Sugano S., White B., Walker D., Woodward J.R., Winckler T.,
RA Tanaka Y., Shaulsky G., Schleicher W., Weinstock G.M., Rosenthal A.,
RA Cox E.C., Chisholm R.L., Gibbs R.A., Loomis W.F., Platzer M.,
RA Kay R.R., Williams J.G., Dear P.H., Noegel A.A., Barrell B.G.,
RA Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum."
RT Nature 435:43-57(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL, AA101000207; EBL6738.1; -; Genomic DNA.
SQ SEQUENCE 532 AA; 60715 MW; BAF75AB8CD7C1D5C CRC64;

```


QY 191 -----PLSEKIAELKEKIV 204
Db 274 DQEBASDARKDQMEVSDLGKSVL 297

RESULT 12
ID Q3V0J4_MOUSE PRELIMINARY; PRT; 497 AA.
AC Q3V0J4;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Adult male testis cDNA, RIKEN full-length enriched library,
DE clone:4930572N15 product:hypothetical Ankyrin/Ankyrin repeat
DE profile/Ankyrin repeat region circular profile containing protein,
DE full insert sequence.
GN Name=4930564N15RIK;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
NP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
[2]
NP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells S., Kodzius R., Shimokawa K.,
RA Raj B., Bremer S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Attalaiya R.N., Bailey T.L.,
RA Banasi M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashizaki Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummelbeck J., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jaki T., Kanapin A., Katoh M., Kawasawa Y., Kelsio J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Moutereau-Tahar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Sempke C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiyama K., Sultana R., Takenaka Y., Taki K.,
RA Tamura J., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yang K.,
RA Yamashita H., Zdobych E., Zhu S., Zimmer A., Hide M., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Watecki J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Aizawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima K., Kondo S., Kono H., Nakano K., Niimura N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami T., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).

EN [3]
NP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC PubMed=16141073; DOI=10.1126/science.1112009;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells S., Kodzius R., Shimokawa K.,
RA Raj B., Bremer S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Attalaiya R.N., Bailey T.L.,
RA Banasi M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashizaki Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummelbeck J., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jaki T., Kanapin A., Katoh M., Kawasawa Y., Kelsio J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Moutereau-Tahar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Sempke C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugiyama K., Sultana R., Takenaka Y., Taki K.,
RA Tamura J., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yang K.,
RA Yamashita H., Zdobych E., Zhu S., Zimmer A., Hide M., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Watecki J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Aizawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima K., Kondo S., Kono H., Nakano K., Niimura N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami T., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[6]
NP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).
 [17]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitanai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubara M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1175-1177(2000).
 [8]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
 RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
 RA Kawai J., Kojima M., Kono H., Murata M., Nakamura M., Nimomiya N.,
 RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
 RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watanaki A.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC EMBL; AK130099; BA021510.1; -; mRNA.
 DR MGI; MGI:1922555; 4930564M15Rik.
 DR GO; GO:0003700; F:transcription factor activity; RCA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; RCA.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; ANK; 3.
 DR PRINTS; PRO1415; ANKYRIN.
 DR SMART; SM00248; ANK; 3.
 DR PROSITE; PSS0297; ANK_REPEAT_1.
 DR PROSITE; PSS0088; ANK_REPEAT_2.
 KW ANK repeat; Hypothetical protein; Repeat.
 KW SEQUENCE 497 AA; 56737 MW; CA87BC48C540354 CRC64;
 SQ
 Query Match 8.0%; Score 96; DB 2; Length 497;
 Best Local Similarity 21.1%; Pred. No. 5.3;
 Matches 64; Conservative 38; Mismatches 89; Indels 112; Gaps 14;
 QY 4 QPRLHNSRCATPPRG-----DPCGTEEA-----ID 29
 DB 3 RPSRRRSKASTPPRSHTTPRRTPGSDSRRRPRTKEOPRPSVGGTRQAEHDLKVSFNS 62
 QY 30 QASFTTSMEDTQVVKGSSPLGAGAEPPAAGPQLPSWLQ-----PER-----C 75
 DB 63 SSOYSISELSMKVIANYSELPAASVGNV-----WLRFCVNDERKEIIVDDKGFT 113
 QY 76 AVQCQAQ-----CHAVLADSVHLAMDLSRSLG-----AVFSRTNNVLEAPLVGREG 125
 DB 114 AHPRAQKQGLSCGLKVLIEYKYKPVLDLPTNKGQRPRLVTHKNNKSDILPCIDYLLKGA 173
 QY 126 SLKGSYTNLLFCGSGCGIPVGFHLYSTHAALAAALR-----GHFCLSS 166
 DB 174 AINSQTN-----GSTPL-----HLASCNGLGCIKTLVQSGANVHARDATGFKPINTYRLW 225
 QY 167 DKWVC-----YLLKTAIVINASEMDIQNVPLSEKIALKEKI-VLTHNRLKSLMK---I 216
 DB 226 NHRTCAFLKDVNMKDKVLAQEM-----EKLTKLKEKLTILEYHYHVLVEYQKHQI 277
 QY 217 LSE 219
 DB 278 LRE 280

ID Q04626 LEIMA PRELIMINARY; PRT; 1556 AA.
 AC Q04626;
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE Hypothetical protein.
 GN ORFNames=LmjF31.2350;
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friledlin;
 RX PubMed=16020728; DOI=10.1126/science.1112680;
 RA Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
 RA Herriman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
 RA Apostolon Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M.,
 RA Blanchettin G., Borzym K., Botne G., Brusch C.V., Collins M.,
 RA Cadag E., Clariot L., Clayton C., Coulson R.M.R., Cronin A.,
 RA Cruz A.K., Davies R.M., De Gaudenzi J., Dobson D.E., Dueterhoeft A.,
 RA Fazellina G., Fosker N., Fraech A.C., Fraser A., Fuchs M., Gabel C.,
 RA Goble A., Goffeau A., Harris D., Hertz-Fowler C., Hilbert H., Horn D.,
 RA Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
 RA Lord A., Louie T., Marra M., Masny D., Matthews K., Michaeli S.,
 RA Motttram J.C., Mueller-Auer S., Munden H., Nelson S., Norbertczak H.,
 RA Oliver K., O'neil S., Pentony M., Pohl T.M., Price C., Purnelle B.,
 RA Quail M.A., Rabinowitsch E., Reinhardt R., Rieger M., Rinta J.,
 RA Robben J., Robertson L., Ruiz J.C., Rutler S., Saunders D.,
 RA Schaefer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
 RA Shin H., Sivam D., Squares R., Squares S., Tostato V., Vogt C.,
 RA Volckreut G., Wambuth D.F., Warren T., Wedler H., Woodward J., Zhou S.,
 RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.G.,
 RA Wyler P.J.;
 RT "The genome of the kinetoplastid parasite, Leishmania major";
 RL Science 309:436-442(2005).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC EMBL; C7005268; CAJ08424.1; -; Genomic_DNA.
 KW Hypothetical protein.
 KW SEQUENCE 1556 AA; 170195 MW; 86F77CDDF8D9CA59 CRC64;
 SQ
 Query Match 7.8%; Score 94; DB 2; Length 1556;
 Best Local Similarity 24.3%; Pred. No. 35;
 Matches 45; Conservative 22; Mismatches 50; Indels 68; Gaps 9;
 QY 22 GGTGRLDQASFTTSE-----WDTQVYKGS-----SPLGPAAGAEPPAAGP 65
 DB 833 GGGEIMQTDAGFSEIVESEPIAFLYDDVHRAQDVQCKRAREBPDVGAGGEEDAAG-- 890
 QY 66 LPSWLDPERCAVFOCAQCHAVLADSV-----HLAMDLSRSLGAVVFSRTNNVLEAP 118
 DB 891 -----APRAIGE--QDFGLNDVTAADPSTYSHMLCRVDELIRTKRLMLILLDLVLAQAP 941
 QY 119 FLVIGIEGLKGSYTNLLFCGSGCGIPVGFHLYSTHAALAAALRGHFLCSSDKMVC---YLL 174
 DB 942 VL-----TRY-----VMYHLK-----HFLCSADKASCVCIGMWLL 971
 QY 175 KTKAI 179
 DB 972 RNLAIV 976
 RESULT 14
 Q05G3_ORYSA
 ID Q05G3_ORYSA PRELIMINARY; PRT; 702 AA.
 AC Q05G3;
 DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
 DT 24-JAN-2006, sequence version 1.
 DT 07-FEB-2006, entry version 2.
 DE Retrotransposon protein, putative, unclassified.
 GN ORFNames=LOC_Os12g24210;

OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BFP clade;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 ON NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Wing R.A., McCombie W.A., Ouyang S.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: DP000011; ABA97657.1; -; Genomic DNA.
 SQ SEQUENCE 702 AA; 76586 MW; 5E958D3E696FF434 CRC64;

Query Match 7.8%; Score 93.5; DB 2; Length 702;
 Best Local Similarity 26.2%; Pred. No. 14;
 Matches 67; Conservative 29; Mismatches 69; Indels 91; Gaps 18;

QY 22 GGERAIDQASFTTSMEMDQOVYKSSPLGPAIGAEPPAGPOLP---SW--LOPERC 75
 DB 335 GGTSSASPAASVAST-----DVALATGSKREATPSG-PVSDPAGAGPPAAVLTWELQVEMR 388
 QY 76 AVFQC-----AQCHAVLADSVHLAW-----DLS--RSLGAV----- 104
 DB 389 RLKAGARISGRIRAEARSEALANT-HADWVRELAEREDLTKRELVAVNERQKGL 447
 QY 105 --VFSRVTVNVLEAPFLVIGESLKGSTYNLL--FCGSGCI-----PVGFHYSTHA 153
 DB 448 EDMSEIGNN-----LSEIRGSLR-VTYTGLHQLAGCGIKSTIPVNPDEFSLNSLA 499
 QY 154 ALAALRG-----HFCLSDKM-----VCYLKTRAIIVNAS--EMDIONV----- 190
 DB 500 ELATWGEIIPSKAARIAETSTNGITYGACHVL--ACVRLSRPELDRLRIDQASDPT 556
 QY 191 --PLSEKIAELKEKIV 204
 DB 557 REEVMKVGDLGESVL 572

RESULT 15
 ID Q8S611 ORYSA PRELIMINARY; PRT; 361 AA.
 AC Q8S611.
 DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2002, sequence version 1.
 DT 21-FEB-2006, entry version 13.
 DE Putative gypsy-type retrotransposon protein.
 GN Name=OSJNBa0096E22.5;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BFP clade;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 ON NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Sasaki C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: AC099400; AAL91599.1; -; Genomic DNA.
 DR Gramene: Q8S611; -;
 SQ SEQUENCE 361 AA; 38805 MW; F883BB3E888FF45 CRC64;

Query Match 7.7%; Score 92; DB 2; Length 361;
 Best Local Similarity 25.3%; Pred. No. 8.4;
 Matches 61; Conservative 28; Mismatches 92; Indels 60; Gaps 14;

QY 2 AAQPLRRHSRCATPPRGDFCGTERAIDQASFTTSMEMDQOVYKSSPLGPAIGAEPPA 61
 DB 49 APTPLPFRRAVRAKAAQDGGGTSSASPAVASV-----DVVVVPGSKREATPSG-PASDPV 102
 QY 62 AGPOLP-----SW--LOPERCAVQCAQCHAVLADSVHLAWDLRSRGAVFSRVTVNVVL 115
 DB 103 AGRGSPAAVLSTWELQVEMGRILAEAGA--RVIGREIAEARGLBHRM-----SEIGNN--- 152
 QY 116 EAPFLVIGESLKGSTYNLL--FCGSGCI-----PVGFHYSTHAATAAL-----RGH 161
 DB 153 ---LSEIRGSLR-VTYTGLHQLAGCGIKSTIPANPDEFSLTSLSAEIQAAMEEIPSKH 207
 QY 162 FCLSDKM-----VCYLKTRAIIVNASEMDIONV-----PLSEKIAELKEKI 203
 DB 208 AARIGEMSVRIYITGACHILACVRLAH-PELDRLRIDQGEASDARKDVMEEVGDGKSV 266
 QY 204 V 204
 DB 267 L 267

Search completed: May 18, 2006, 15:40:51
 Job time : 302 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006, Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2006, 15:41:07 ; Search time 50 Seconds
(without alignments)
400.890 Million cell updates/sec

Title: US-09-942-052A-728

Perfect score: 1198

Sequence: 1 MAOPLRHRRCATPPRGDF.....LKSIMKILSEVTPDQSKPEN 229

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /EMC_Celerra_SIDS3/prodata/2/1aa/5 COMB pep:*
- 2: /EMC_Celerra_SIDS3/prodata/2/1aa/6 COMB pep:*
- 3: /EMC_Celerra_SIDS3/prodata/2/1aa/7 COMB pep:*
- 4: /EMC_Celerra_SIDS3/prodata/2/1aa/H COMB pep:*
- 5: /EMC_Celerra_SIDS3/prodata/2/1aa/PCTUS COMB pep:*
- 6: /EMC_Celerra_SIDS3/prodata/2/1aa/RE COMB pep:*
- 7: /EMC_Celerra_SIDS3/prodata/2/1aa/backfile1 pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	7.2	158	US-09-134-000C-6596	Sequence 6596, Ap
2	83	6.9	457	5268463-7	Patent No. 5268463
3	81.5	6.8	456	5432081-7	Patent No. 5432081
4	81.5	6.8	457	US-08-882-704A-6	Sequence 6, Appli
5	81.5	6.8	457	US-09-151-957-6	Sequence 6, Appli
6	80.5	6.7	1711	US-10-195-518-6	Sequence 6, Appli
7	80.5	6.7	1711	US-08-342-930-2	Sequence 2, Appli
8	78.5	6.6	325	US-09-018-576-3	Sequence 3, Appli
9	78.5	6.6	325	US-09-018-576-3	Sequence 12, Appli
10	78.5	6.6	325	US-08-248-137-3	Sequence 3, Appli
11	78.5	6.6	325	US-09-248-137-12	Sequence 12, Appli
12	78.5	6.6	452	US-09-771-161A-237	Sequence 237, Appli
13	78	6.5	190	US-09-252-991A-19049	Sequence 19049, A
14	77	6.4	607	US-09-537-682-1	Sequence 1, Appli
15	77	6.4	719	US-08-520-933-3	Sequence 3, Appli
16	77	6.4	719	US-09-285-040-3	Sequence 3, Appli
17	77	6.4	738	5262177-2	Patent No. 5262177
18	77	6.4	776	US-09-949-016-10829	Sequence 10829, A
19	76.5	6.4	282	US-09-845-713A-2	Sequence 2, Appli
20	76	6.3	351	US-09-902-540-13214	Sequence 13214, A
21	76	6.3	1001	US-10-104-047-2283	Sequence 2283, Ap
22	75.5	6.3	310	US-08-129-456A-36	Sequence 36, Appli
23	75.5	6.3	310	US-08-705-868-3	Sequence 3, Appli
24	75.5	6.3	310	US-09-123-615-3	Sequence 3, Appli
25	75.5	6.3	310	US-08-360-821B-35	Sequence 35, Appli
26	75.5	6.3	310	US-09-851-026-35	Sequence 35, Appli

27	75.5	6.3	1449	US-08-840-062-6	Sequence 6, Appli
28	75.5	6.3	1723	US-09-194-612A-31	Sequence 31, Appli
29	75.5	6.3	4545	US-08-804-227C-14	Sequence 14, Appli
30	75.5	6.3	4550	US-08-804-227C-8	Sequence 8, Appli
31	75.5	6.3	4550	US-08-804-198-2	Sequence 2, Appli
32	75	6.3	78	US-09-270-767-43496	Sequence 43496, A
33	75	6.3	458	US-09-800-170-3	Sequence 3, Appli
34	75	6.3	717	5262177-5	Patent No. 5262177
35	75	6.3	784	US-09-724-586A-1	Sequence 1, Appli
36	75	6.3	784	US-09-600-823-1	Sequence 1, Appli
37	75	6.3	784	US-09-724-666-1	Sequence 1, Appli
38	75	6.3	1088	US-09-920-804-2	Sequence 2, Appli
39	75	6.3	1088	US-09-920-804-4	Sequence 4, Appli
40	75	6.3	1129	US-09-734-674-2	Sequence 2, Appli
41	75	6.3	1129	US-10-274-990-2	Sequence 2, Appli
42	75	6.3	2227	US-08-475-886-4	Sequence 4, Appli
43	75	6.3	2227	US-09-653-499-4	Sequence 4, Appli
44	75	6.3	2227	US-10-135-988-4	Sequence 4, Appli
45	74.5	6.2	211	US-09-949-016-11409	Sequence 11409, A

ALIGNMENTS

```
RESULT 1
US-09-134-000C-6596
; Sequence 6596, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6596
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6596

Query Match 7.2%; Score 86; DB 2; Length 158;
Best Local Similarity 26.5%; Pred. No. 0.19;
Matches 40; Conservative 27; Mismatches 54; Indels 30; Gaps 9;

QY 87 LADSVHAMDLSRLGAVV---FSRYNNVYLA--PLVIGRSGK-----STVN 133
   |::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 17 LTYIYIYASQPISTGISTINGLSKPTGNLRAFFFIINDFGDPGRKFSYNYLKTIN 76
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 134 LIFCGSGIFVGFHLYSTHAAL--RGHFLSSDKWVCYLTKTAYVASEMDIONV 190
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 77 YVCSMSLPSDQNIETPAIALLGIRNH-----KATC---TKLYIISLNGRENK 128
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 191 PLSE---KIAELKEIVLTHNRKSLMKIL 217
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 129 PISKTVRKLSLSYK-DLYQEMK-LTKII 157
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 2
5268463-7
; Patent No. 5268463
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: PLANT PROMOTER a-GLUCURONIDASE GENE
; CONSTRUCT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,976
; FILING DATE: 08-DEC-1989
; PRIOR APPLICATION DATA:
```


CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/882,704
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 6429292tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 190106.404
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-151-957-6
Query Match 6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 3.2;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;
QY 95 WDLRSIGAVFVRVTVN-----VLEA-----PFLVGIESLKGSTYNL-FCGSCG 141
DB 317 WSLPVALVALAISIGGVMTVMWALEADTVGEYLTGVR--IEGLTSLFSFTKCG 374
QY 142 IPVGFHLYSTHALALRGHFC--LSSDKVVCYLTKTAIVNSENDIONV-----PLSE 194
DB 375 QAIG---GSIPAFILGSGYIANQVQPEYIMGIRTSIALVPCGFMLAFVITWYPLTD 431
QY 195 KIAELKEKIVLTNRRLKSLMKILSEVT 221
DB 432 K--KFKEIIVEIDNRKKVQQQLISDIT 456
RESULT 6
US-10-195-518-6
Sequence 6, Application US/10195518
Patent No. 6998229
GENERAL INFORMATION:
APPLICANT: Jefferson, Richard A.
Leader, Michael
TITLE OF INVENTION: GLUTCRONIDE REPRESSORS AND USES THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/195,518
FILING DATE: 16-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/151,957
FILING DATE: 11-Sep-1998
APPLICATION NUMBER: US 08/882,704
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 6998229tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 190106.404
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-195-518-6
Query Match 6.8%; Score 81.5; DB 2; Length 457;
Best Local Similarity 25.2%; Pred. No. 3.2;
Matches 37; Conservative 30; Mismatches 53; Indels 27; Gaps 8;
QY 95 WDLRSIGAVFVRVTVN-----VLEA-----PFLVGIESLKGSTYNL-FCGSCG 141
DB 317 WSLPVALVALAISIGGVMTVMWALEADTVGEYLTGVR--IEGLTSLFSFTKCG 374
QY 142 IPVGFHLYSTHALALRGHFC--LSSDKVVCYLTKTAIVNSENDIONV-----PLSE 194
DB 375 QAIG---GSIPAFILGSGYIANQVQPEYIMGIRTSIALVPCGFMLAFVITWYPLTD 431
QY 195 KIAELKEKIVLTNRRLKSLMKILSEVT 221
DB 432 K--KFKEIIVEIDNRKKVQQQLISDIT 456
RESULT 7
US-08-342-930-2
Sequence 2, Application US/08342930
Patent No. 5821084
GENERAL INFORMATION:
APPLICANT: OLIMSTED, ELIZABETH A.
APPLICANT: MAURO, LAURA J.
APPLICANT: DAVIS, ALAN R.
APPLICANT: DIXON, JACK E.
TITLE OF INVENTION: OSTROBLAST-TESTICULAR PROTEIN TYROSINE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/342,930
FILING DATE: 21-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KONSKI, ANTOINETTE F.
REGISTRATION NUMBER: 34,202
REFERENCE/DOCKET NUMBER: 20344-20975.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-342-930-2
Query Match 6.7%; Score 80.5; DB 1; Length 1711;

PRIOR APPLICATION NUMBER: 135619
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 273
SOFTWARE: Patentin version 3.0
SEQ ID NO 237
LENGTH: 452
TYPE: PRT
ORGANISM: Homo sapiens
US-09-771-161A-237

Query Match 6.6%; Score 78.5; DB 2; Length 452;
Best Local Similarity 23.2%; Pred. No. 6.9;
Matches 56; Conservative 24; Mismatches 78; Indels 83; Gaps 12;

QY 22 GGTGTAIDQASFTSMEMDTQVVKSSPLGPAAG-----AEEPAG 63
DB 43 GEPNOLREIKALQEME-DNQYVQKAVFPHGGVLAEEFMLSDEAVVRHAQRPLO 101
QY 64 POLPSWLOPERCAVFCQACGAVLADSVHLAMDLSRSLGAVFSRVTVNNV---LEAPF 119
DB 102 AOVKSTLOMLKKV---AFCHA-----NNIVHDLKPAUL 133
QY 120 LVGIEGSLKSTYNLLFCGSCGIPVGFHYSTHAALALRGHFCLSDD----- 167
DB 134 LISASGOLKIADPGLARVFS---PDGSRLY-THQ--VATRSVGCIMGELINGSPLEPGKN 187
QY 168 --KMCVYLKTKAIIVNASEM-DIQNVPLSEKIAELKEKIVLTHNRLKSLMKILSEVTPDQ 224
DB 188 DIEQLCYVNIIGTPEPQVPELLELPDYNKIS-FKEQVWP-----LEEVLDPV 236
QY 225 S 225
DB 237 S 237

RESULT 13
US-09-252-991A-19049
Sequence 19049, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19049
LENGTH: 190
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19049

Query Match 6.5%; Score 78; DB 2; Length 190;
Best Local Similarity 32.9%; Pred. No. 2.1;
Matches 24; Conservative 6; Mismatches 29; Indels 14; Gaps 3;
QY 4 OPLHRSRCATPPRGDFCGTERTAIQASFTSMEMDTQVVKSSPLGPAAGLEEPAG 63
DB 94 RPCRHRRCYCPAAGGTPG---RAGDPAAS-----VARSIGPOLPAAPGRPRGHC- 139
QY 64 POLPSWLOPERCA 76
DB 140 PAPPRWRRRRCA 152

RESULT 14
US-09-537-682-1
Sequence 1, Application US/09537682

Patent No. 6303357
GENERAL INFORMATION:
APPLICANT: TAKEUCHI, Kenichi
APPLICANT: KOIDE, Yoshinao
APPLICANT: NAKANISHI, Yuji
APPLICANT: SUZUKI, Satoru
TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE, RECOMBINANT
TITLE OF INVENTION: DNA, AND METHOD FOR PRODUCING MODIFIED
TITLE OF INVENTION: L-ALPHA-GLYCEROPHOSPHATE OXIDASE GENE
FILE REFERENCE: A20-123814C/KI
CURRENT APPLICATION NUMBER: US/09/537,682
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 607
TYPE: PRT
ORGANISM: Enterococcus faecium No. 6303357 7044
US-09-537-682-1

Query Match 6.4%; Score 77; DB 2; Length 607;
Best Local Similarity 23.2%; Pred. No. 16;
Matches 49; Conservative 30; Mismatches 76; Indels 56; Gaps 10;

QY 19 DFCGTERAIDQASFTSM-----EMDTQVVKSSPLGP---AGLGAEEPAAGPQ 65
DB 51 DFAEGT-----SSRETKLHVGIRYLTFTFDEVV--ADTVGERAVVQGIAPHIPKDDPM 102
QY 66 LPSWLOPERCAVFCQACGAVLADSVHLAMDLSRSLGAVFSRVTVN-----NVLEAPF 119
DB 103 LPEYDEGATTFNMF-----SVKAMDLYDKLANVTGKVENYTLTPBEVLERBEF 154
QY 120 LVGIEGSLKSTYNLLFCGSCGIPVGF-----HLYSTHAALALRGHFCLSDPKWCYLL 174
DB 155 LK--KEGLKA-----GVYIDFRNNDAFLVDNKKKAEDGAYLVSKKAVGFLY 202
QY 175 KTKAIIVNASEMDIQNVPLSEKIAELKEKIVL 205
DB 203 EGDQIVGVAKARDL---LTDEVIERSKLV 229

RESULT 15
US-08-520-933-3
Sequence 3, Application US/08520933
Patent No. 5981194
GENERAL INFORMATION:
APPLICANT: Jefferies, Wilfred A.
APPLICANT: McGeer, Patrick L.
APPLICANT: Rothenberger, Sylvia
APPLICANT: Food, Michael R.
APPLICANT: Yamada, Tatsuo
APPLICANT: Kennard, Malcolm
TITLE OF INVENTION: Use of p97 and Iron Binding Proteins
TITLE OF INVENTION: as Diagnostic and Therapeutic Agents
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/520,933
FILING DATE: August 31, 1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Shona S. McDiarmid

REGISTRATION NUMBER: 38,798
REFERENCE/DOCKET NUMBER: 7685-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-364-7311
TELEFAX: 416-361-1398
TELEX: 06-23115
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-520-933-3

Query Match 6.4%; Score 77; DB 1; Length 719;
Best Local Similarity 23.7%; Pred. No. 21;
Matches 59; Conservative 25; Mismatches 87; Indels 78; Gaps 15;

QY	17	RGDFCGGTERAIDQASFTSMEW-----DTQVVKSSPILGPAGIGAEPAAGP	64
DB	171	RGDSSG--EGVCDKSPLERYDYSGAFRCIAEGAGDVAFVKHSTVL-----ENTDGK	220
QY	65	QLPSWLTQP-----ERCAYFQCAQCH--AVLADSVHLAMDLSRSLGAVFSRV	109
DB	221	TLPSWGOALISQDFELLCRDGSRADVTETWRQCHLARVPAAHAVVVRADTD--GGLIF-RL	276
QY	110	TNNVVLKAPFLVIGEGS---LKSTY---NLFCGSG--IPVGFHLYST--HALAA	157
DB	277	LN---EGQRLFSHEGSSFGMFSEAYGQKDLFKDSTSELVPIATQTYEAMLGHEYLAH	332
QY	158	LRGHFCLSDKWCYL---LKTKAIVNASEMDIQNVPLSEKIAELKEKIVLTNNRLKSL	213
DB	333	MKGILC-DPNRLPPYLRLWCYLTSTPEIQKCGM-----AVALFRQRLLKPE	375
QY	214	MKILSEVTP	222
DB	376	IQCVSAKSP	384

Search completed: May 18, 2006, 15:42:30
Job time : 51 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocolloration Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2006, 15:52:57 / Search time 182 Seconds
(without alignments)
582.836 Million cell updates/sec

Title: US-09-942-052a-728

Perfect score: 1198

Sequence: 1 MAOPLRHRRCATPPRGDF.....LKSIMKILSEVTPDQSKPEN 229

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1198	100.0	229	US-09-942-052-728	Sequence 728, App
2	1198	100.0	229	US-09-942-052-729	Sequence 729, App
3	1198	100.0	229	US-09-942-052-730	Sequence 730, App
4	1198	100.0	229	US-09-942-052-731	Sequence 731, App
5	138.5	11.6	164	US-09-942-052-707	Sequence 707, App
6	138	11.5	29	US-09-942-052-704	Sequence 704, App
7	113	9.4	23	US-09-890-688-82	Sequence 82, App
8	113	9.4	23	US-10-408-765A-969	Sequence 969, App
9	95.5	8.0	708	US-10-437-963-195427	Sequence 195427, App
10	92	7.7	361	US-10-369-493-18389	Sequence 18389, App
11	92	7.7	695	US-10-369-493-18389	Sequence 18389, App
12	89	7.4	664	US-09-780-525-2	Sequence 2, App11
13	88.5	7.4	435	US-10-437-963-186569	Sequence 186569, App
14	87.5	7.3	848	US-10-437-963-155606	Sequence 155606, App
15	86.5	7.2	669	US-10-282-122A-51802	Sequence 51802, App
16	86	7.2	584	US-11-096-568A-12723	Sequence 12723, App
17	86	7.2	619	US-11-096-568A-12722	Sequence 12722, App
18	86	7.2	644	US-11-096-568A-12721	Sequence 12721, App
19	85	7.1	337	US-10-156-761-10954	Sequence 10954, App
20	84.5	7.1	295	US-10-363-829-358	Sequence 358, App
21	84	7.0	295	US-10-424-599-223506	Sequence 223506, App
22	84	7.0	1902	US-10-437-963-139559	Sequence 139559, App
23	83.5	7.0	888	US-10-437-963-147897	Sequence 147897, App
24	83	6.9	663	US-10-282-122A-51673	Sequence 51673, App
25	82.5	6.9	124	US-10-425-115-216789	Sequence 216789, App
26	82.5	6.9	278	US-10-425-114-42310	Sequence 42310, App
27	82.5	6.9	325	US-10-663-896-2	Sequence 2, App11

28	82.5	6.9	493	US-10-213-974-36	Sequence 36, App1
29	82.5	6.9	502	US-10-424-599-226730	Sequence 226730, App
30	82.5	6.9	1043	US-10-369-493-9793	Sequence 9793, App
31	81.5	6.8	278	US-10-425-114-66220	Sequence 66220, App
32	81.5	6.8	457	US-10-195-518-6	Sequence 518, App
33	81.5	6.8	701	US-10-450-763-115342	Sequence 115342, App
34	81.5	6.8	1969	US-10-437-963-115342	Sequence 115342, App
35	81.5	6.8	3067	US-10-631-467-1618	Sequence 1618, App
36	81.5	6.8	229	US-10-425-115-281194	Sequence 281194, App
37	81	6.8	859	US-10-210-281-60	Sequence 60, App1
38	80.5	6.7	282	US-10-450-763-51058	Sequence 51058, App
39	80.5	6.7	1711	US-10-029-345A-38	Sequence 38, App1
40	80.5	6.7	1711	US-11-143-984A-38	Sequence 38, App1
41	80	6.7	250	US-11-079-984A-38	Sequence 9254, App
42	80	6.7	305	US-10-183-116-59	Sequence 59, App1
43	80	6.7	305	US-10-957-135-59	Sequence 59, App1
44	80	6.7	305	US-11-083-611-59	Sequence 59, App1
45	80	6.7	544	US-10-424-599-158084	Sequence 158084, App

ALIGNMENTS

RESULT 1
US-09-942-052-728
Sequence 728, Application US/09942052
Publication No. US20030170626A1
GENERAL INFORMATION:
APPLICANT: Raitano, Arthur B.
APPLICANT: Farris, Mary
APPLICANT: Hubert, Rene S.
APPLICANT: Afar, Daniel
APPLICANT: Ge, Wangmao
APPLICANT: Challita-Eid, Pia M.
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
FILE REFERENCE: 51158-20028.00
CURRENT APPLICATION NUMBER: US/09/942,052
CURRENT FILING DATE: 2001-08-28
PRIORITY APPLICATION NUMBER: 60/228,432
PRIORITY FILING DATE: 2000-08-28
NUMBER OF SEQ ID NOS: 744
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 728
LENGTH: 229
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIPs
OTHER INFORMATION: clone A protein
US-09-942-052-728

Query Match 100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-113;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAOPLRHRRCATPPRGDFCGGTERAIDQASTTSMENDTVYKSSPIGPGGLAEEP	60
DB	1	MAOPLRHRRCATPPRGDFCGGTERAIDQASTTSMENDTVYKSSPIGPGGLAEEP	60
QY	61	AAQPOLPSWLPQRCVAFQCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL	120
DB	61	AAQPOLPSWLPQRCVAFQCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL	120
QY	121	VGIEGSLKSGTYNLLFCGSGGIPVGFHLVSTHAALALRGHFLCLSSDKWVCYLKTKAIV	180
DB	121	VGIEGSLKSGTYNLLFCGSGGIPVGFHLVSTHAALALRGHFLCLSSDKWVCYLKTKAIV	180
QY	181	MAEMDIQWVPLSEKIAELKEKIVLTHNLKSLMKILSEVTPDQSKPEN	229
DB	181	MAEMDIQWVPLSEKIAELKEKIVLTHNLKSLMKILSEVTPDQSKPEN	229

```
RESULT 2
US-09-942-052-729
; Sequence 729, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Fatis, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Aftar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eld, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942.052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 729
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIPs
US-09-942-052-729

Query Match          100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-113;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLHRRCATPPRGDFCGGTERAIDQASFTTSMENDTOVVKSSPLGPAGLGAEEP 60
    |||||
DB 1 MAAPLHRRCATPPRGDFCGGTERAIDQASFTTSMENDTOVVKSSPLGPAGLGAEEP 60

QY 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL 120
    |||||
DB 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL 120

QY 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180
    |||||
DB 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180

QY 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180
    |||||
DB 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180

QY 181 NASEMDIONVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDOSKPEEN 229
    |||||
DB 181 NASEMDIONVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDOSKPEEN 229

RESULT 3
US-09-942-052-730
; Sequence 730, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Fatis, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Aftar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eld, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942.052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 730
; LENGTH: 229
; TYPE: PRT
```

```
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3 protein
; OTHER INFORMATION: sequence
US-09-942-052-730

Query Match          100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-113;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLHRRCATPPRGDFCGGTERAIDQASFTTSMENDTOVVKSSPLGPAGLGAEEP 60
    |||||
DB 1 MAAPLHRRCATPPRGDFCGGTERAIDQASFTTSMENDTOVVKSSPLGPAGLGAEEP 60

QY 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL 120
    |||||
DB 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL 120

QY 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180
    |||||
DB 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180

QY 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180
    |||||
DB 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180

QY 181 NASEMDIONVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDOSKPEEN 229
    |||||
DB 181 NASEMDIONVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDOSKPEEN 229

RESULT 4
US-09-942-052-731
; Sequence 731, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Fatis, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Aftar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eld, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942.052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 731
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-942-052-731

Query Match          100.0%; Score 1198; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-113;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAPLHRRCATPPRGDFCGGTERAIDQASFTTSMENDTOVVKSSPLGPAGLGAEEP 60
    |||||
DB 1 MAAPLHRRCATPPRGDFCGGTERAIDQASFTTSMENDTOVVKSSPLGPAGLGAEEP 60

QY 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL 120
    |||||
DB 61 AAGPOLPSWLQPERCAVFOCAQCHAVLADSVHLAMDLSRLGAVFSRVTNNVLEAPFL 120

QY 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180
    |||||
DB 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180

QY 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180
    |||||
DB 121 VGIEGSLKGSTYNLLFCGSGCIPVGPHLYSTHAALALRGHFCCLSDKWCYLLKTKTAIV 180

QY 181 NASEMDIONVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDOSKPEEN 229
    |||||
DB 181 NASEMDIONVPLSEKIAELKEKIVLTNRLKSLMKILSEVTPDOSKPEEN 229
```

```
RESULT 5
US-09-942-052-707
; Sequence 707, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afari, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Chailita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Splice variant 1
; OTHER INFORMATION: open reading frame 3 peptide sequence
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (44)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (50)..(51)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (67)
; OTHER INFORMATION: Unknown amino acid or stop
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (93)
; OTHER INFORMATION: Unknown amino acid or stop
US-09-942-052-707

Query Match          11.6%; Score 138.5; DB 3; Length 164;
Best Local Similarity 82.1%; Pred. No. 1.3e-05;
Matches 32; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY      162 FCSSDKMVCYLKTKAIVNASEMDIONVPLSEKIAELK 200
Db      2 FLLSS-----YLLKTKAIVNASEMDIONVPLSEKIAELK 35

RESULT 6
US-09-942-052-704
; Sequence 704, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afari, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Chailita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
```

```
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 704
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: 85P1B3 peptide
US-09-942-052-704

Query Match          11.5%; Score 138; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      172 YLLKTKAIVNASEMDIONVPLSEKIAELK 200
Db      1 YLLKTKAIVNASEMDIONVPLSEKIAELK 29

RESULT 7
US-09-890-688-82
; Sequence 82, Application US/09890688
; Publication No. US20030144475A1
; GENERAL INFORMATION:
; APPLICANT: Chikashi EGUCHI
; APPLICANT: Mihoro SAKI
; TITLE OF INVENTION: Human Proteins and cDNAs thereof
; FILE REFERENCE: 2001-1102A/WMC/00653
; CURRENT APPLICATION NUMBER: US/09/890,688
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: JP 11-346863
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 11-34684
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 2000-31062
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: JP 2000-34091
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-34090
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-35829
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-35899
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-71161
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: JP 2000-160851
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-890-688-82

Query Match          9.4%; Score 113; DB 3; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0083;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8;

QY      31 ASFTSMENDTQVKKSSPLGPGAGCAEPAAGPQLPSMLQPRCAVFOCAOCHAVLADS 90
Db      49 ASWSSMSSEDAVY---ADMERAGQ---EBEAAAE-----ERPLVFLSCGRRPLGDS 95

QY      91 VHLAMPISR-SLGAIVFSRVTNNVLEAPFLVIGESLGSFTYNLLFCGSCGIPVGFHHY 149
Db      96 --LSWVASOEDNTICILRCVSCNVSVDKQKLSKREKNGCVLETICGAGCSINLGIYVR 153

QY      150 STHAALALRGHFCSSDKMVCYL--KTKAIVNASEMDIONVPLSEKIAELKIVLTH 207
Db      154 CTPKNLDYKRDLFCLSVFAIESYVLGSSSEKQIV--SEDKELFNL---ESRVEIEKSLTQME 209
```

Qy	208	NRLKSLMKILSE	219
	:	:	
Db	210	DVLKALQMKLWE	221

RESULT 8
US-10-408-765A-969

```

Sequence 969, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Paly, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088, 465
CURRENT APPLICATION NUMBER: US/10/408, 765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 969
LENGTH: 233
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-969

```

Query Match 9.4%; Score 113; DB 4; Length 233;
Best Local Similarity 27.1%; Pred. No. 0.0083;
Matches 52; Conservative 30; Mismatches 88; Indels 22; Gaps 8.

[illegible]

RESULT 9
US-10-437-963-195427

```

: Sequence 195427, Application US/10437963
: Publication NO. US20040123343A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Wu, Wei
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Barbazuk, Brad
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53221)B
: CURRENT APPLICATION NUMBER: US/10/437,963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ. ID NOS: 204966
: SEQ. ID NO 195427
: LENGTH: 708

```

```

1  TYPE: PRT
2  /
3  ORGANISM: Oryza sativa
4  /
5  FEATURE:
6  /
7  OTHER INFORMATION: Clone ID: PAT_MRT4530_91377C.1.pep
8  /
9  US-10-437-963-195427

```

Query Match	8.0%	Score 95.5	DB 4	Length 708
Best Local Similarity	24.5%	Pred. No. 2.3		
Matches 71	Conservative 32	Mismatches 86	Indels 101	Gaps 17

[illegible]

RESULT 10

```

US-10-437-963-154548
; Sequence 154548, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154548
; LENGTH: 361
; TYPE: PRP
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54396C.1.pep
US-10-437-963-154548

```

Query Match	7.7%	Score 92	DB 4	Length 361
Best Local Similarity	25.3%	Pred. NC	2.1	
Matches 61, Conservative	28	Mismatches	92	Indels 60, Gaps 14

```

QY      2  AAOPLHRSKCAIPPRGDCGGTEBAIDQASTTSMEDTVQVGGSPBLAGAGAEPA  61
Db      49  APPLPRRAVRAKAAQDSGGTSSASPAVAST-----DVAVVGSRREAPSG--PASDPV  102
QY      62  AGPQLP-----SW--LQPERCAVFOCAQCCHAVLADSVHLAMDLSRSLGAVFPSRYTNNVL  115
Db      103  AGKSPPAVLISWEIQLVEMGRLLKGA--RVIGREIARGLIHRM-----SELGN-----  155
QY      116  EAPFLVIGESLKGSTYNLL--FCGSCGI-----PGFHLVSTHALAAL-----RGH  161
Db      153  ----LSTKRSLLR-VYTTGLHQLAGKCGIKSTIIPANPPEFSLTSLAEIILAAAMEIPISKH  207

```


Qy 162 FCLSSDKM-----VCYLKTKAIVNASEMDIONV-----PLSPKIAELKEKI 203
Db 208 AARIGSEMSRIYAGCHILACVRLAH-PELDREILDOGEASDARKDVMEVEVDLGSV 266
- Qy 204 V 204
Db 267 L 267

RESULT 11
US-10-369-493-18389
; Sequence 18389, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18389
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Lactococcus lactis
US-10-369-493-18389

Query Match 7.7%; Score 92; DB 4; Length 695;
Best Local Similarity 25.3%; Pred. No. 5.2;
Matches 38; Conservative 28; Mismatches 50; Indels 34; Gaps 7;
Qy 101 LGAVFSRVNNVLEAPLVGIGSLKGSYTNLLFCGSCGIPVGFHLYSTHALALRG 160
Db 101 LGRIIF-----FVSGTFPFEGAGELKSRKPMAMMILTMGITVAY-AVSYYATIMSLG 153
Qy 161 HF-----CLSSDKMVCYLKTKAIVNASE--MDI-----QNVPLSE-K 195
Db 154 HMGNNFPELATLIVIMLIGLITMKAIMGAGDLKODASIVPKKAHLKSGKDVLSLK 213
Qy 196 IAEIKERIVLTNNLKSIMKLISEVTPDOS 225
Db 214 VGDL--LLVKENEKIPADGILSLSEAVDES 241

RESULT 12
US-09-780-525-2
; Sequence 2, Application US/09780525
; Patent No. US20020004223A1
; GENERAL INFORMATION:
; APPLICANT: Bin-Bing Zhou
; APPLICANT: Yuan Zhu
; APPLICANT: Priya Chaturvedi
; APPLICANT: Mark R. Hurle
; APPLICANT: Xiaolong Li
; TITLE OF INVENTION: PHAR, A NEW RING FINGER PROTEIN
; FILE REFERENCE: GP-70668-C1
; CURRENT APPLICATION NUMBER: US/09/780,525
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 09/456,876
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: HOMO SAPIENS

US-09-780-525-2
Query Match 7.4%; Score 89; DB 3; Length 664;
Best Local Similarity 23.0%; Pred. No. 9.9;
Matches 52; Conservative 26; Mismatches 74; Indels 74; Gaps 11;
Qy 8 HRSRCATPPRGDFC-----GTERAIDQA-----SFTSMEMDTQVKGSSPL----- 50
Db 434 YRRQAQPPH---CPABEGEPGAPQALGDAPSTSVSLTTAVQDVVCLQSHALCTCCFQ 490
Qy 51 ---GPAGLGAEBPAGPQLPSWLOPERCAVFOCAQCHAVLADSVHLAMDLSR----- 99
Db 491 PMPDRAERERQDPVAPQ-----QCAVC---LQPCHLVWGCTRTGCGYCLA 534
Qy 100 -----SLGAVFSRYTNVLEAPLVGIGSLKGSY--NLFCGSCGIVGFHLYSTHA 153
Db 535 PCELNLDGKCLDGVNLNNSYESDILKNVLYAT-RGLTWKML-----TES 578
Qy 154 ALAALRGHPLCS-----SDKVCYLKTKAIVNASEMDIONVPLSE 194
Db 579 LVALLQGVFLSDRYVGTIVLCYCCGLRSFRELITYRONIPASE 624

RESULT 13
US-10-437-963-186569
; Sequence 186569, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 186569
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83357C.1.pep
US-10-437-963-186569

Query Match 7.4%; Score 88.5; DB 4; Length 435;
Best Local Similarity 22.9%; Pred. No. 6.2;
Matches 50; Conservative 27; Mismatches 82; Indels 59; Gaps 12;
Qy 2 AAOPL--RH--RSRCATPPRGDFCGTERAIDQASFTSMEMDTQVKGSSPLGAGIGA 57
Db 41 APEPLSCRHRHRCAAVD-----GGAGRTERPSPAPQ-----RESPGSLGAL 88
Qy 58 EBPAGPQLPSWLOPERCAVFOCAQCHA--VLADSVHLADLSRSLGAVVFSRYTNVVL 115
Db 89 EDPSPQGVPLILP-----LCRCYAKKICSEYVVRTDLVNH-----LNSVAIS 135
Qy 116 EAPLVGIGSLKGSYTNLLFCGSCGIPVGFHLYSTHALALRGHFLCSSDKMVCYLK 175
Db 136 EGPF-----SMRKARFLGSAS--AFSVQTEWP-----CATTDKCY-LQ 173
Qy 176 TKAIVNASEMDION-----VPLSEKIAELKEKIVLTH 207
Db 174 NGFSGITDEBQSLYNPLFYSKELLPDKEMSIPIH 211

RESULT 14
US-10-437-963-155606

```
/ Sequence 155606, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Bouharov, Andrey A.
/ APPLICANT: Barabuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 155606
/ LENGTH: 848
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(848)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_55353C.1.pep
US-10-437-963-155606
```

```
Query Match
Best Local Similarity 7.3%; Score 87.5; DB 4; Length 848;
Matches 65; Conservative 23; Mismatches 109; Indels 97; Gaps 13;
```

```
QY 16 PRGFGCGTERAIDQASFTTSMEMDQVVGSSPLGAGAEPP----- 60
DB 85 PRGRLDGTGGVAGQRTASCHLTRDLPFLKSGGLGGRGKRNPPPLGSGNGBERHL 144
QY 61 AAGPOLPSWLOPERCAVFCQACQ-----AVLADSVHLAMDLSR----- 99
DB 145 TLBPDEVGARGPRRLCLPQASGHDTPGPISPTSLVINKRRIRIDEVATMDISEANEG 204
QY 100 --SLGAVV-FSRVTNNV-----VLEAPPLVIGIESLKGSTYVLLFCG----- 138
DB 205 YVSGSVIEMSRQKAKARAGVRAQSPACKGIPYGVKGSV--LLEFARMGEGLCHMFTT 262
QY 139 -----SGGIPV-----GPHLYSTH---AALALRGHFCLSSDKMVCYL 173
DB 263 GSKAMECGVKKCVGWFWLAWDELGPRSGYQFGLNHRRPLVKAVLDGHLCEEDAVNCLP 322
QY 174 LKTKAIVNASEMDIQNVPLSEKIAELKEKIVLTNRRLKSLMKILSEVTPDQSKP 227
DB 323 KLRRTVGSASE-----AKFAVKPAVKQEKIIRS-VKVLISVS-DLSLP 363
```

RESULT 15

US-10-282-122A-51802

/ Sequence 51802, Application US/10282122A

/ Publication No. US20040029129A1

/ GENERAL INFORMATION:

/ APPLICANT: Wang, Liangsu

/ APPLICANT: Zamudio, Carlos

/ APPLICANT: Malone, Cheryl

/ APPLICANT: Haebelbeck, Robert

/ APPLICANT: Ohlsen, Kari

/ APPLICANT: Zykend, Judith

/ APPLICANT: Wall, Daniel

/ APPLICANT: Trawick, John

/ APPLICANT: Carr, Grant

/ APPLICANT: Yamamoto, Robert

/ APPLICANT: Forsyth, R.

/ APPLICANT: Xu, H.

/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

/ FILE REFERENCE: EITRA.034A

```
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 76614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 51802
/ LENGTH: 669
/ TYPE: PRT
/ ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51802
```

```
Query Match
Best Local Similarity 7.2%; Score 86.5; DB 4; Length 669;
Matches 37; Conservative 22; Mismatches 45; Indels 31; Gaps 8;
```

```
QY 96 DLSR---SLGAVVFSRVTNNVLEAPPLVIGIESLKGSTYNL-----LFCGCGIPIYGF 146
DB 366 DISRKKVKGSRVFRNSNDVI--PEIMGVTEETGETTEIEAPITCPYCGSEIYKEGV 422
QY 147 HLYSTHAALALGHFCLSSDKMVCYLTKAIVNASEMDIQNVPLSEKIAE-LXKIV 204
DB 423 HL-----FC--ENTLSCKQMWKSLVHFAISRAMNIBGSEKTAQLERK-- 465
QY 205 LTHNRKLSLMLKILSE 219
DB 466 LNIKSIDLYITKE 480
```

Search completed: May 19, 2006, 15:56:08
Job time : 183 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2006, 15:53:12 ; Search time 10 seconds
(without alignments)
48.917 Million cell updates/sec

Title: US-09-942-052a-728

Perfect score: 1198
Sequence: 1 MAOPLRHRSRCATPPRGDF.....LKSIMKILSEVTPDQSKPEN 229

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 21570 seqs, 2136119 residues

Total number of hits satisfying chosen parameters: 21570

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA_New:*
1: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71.5	6.0	1006	US-10-511-937-2425	Sequence 2425, App
2	67	5.6	388	US-11-312-958-44	Sequence 44, App
3	67	5.6	672	US-10-370-959-152	Sequence 152, App
4	66.5	5.6	345	US-11-181-115-5	Sequence 5, App
5	66.5	5.6	345	US-11-181-115-8	Sequence 8, App
6	66.5	5.6	345	US-11-181-115-12	Sequence 12, App
7	66.5	5.6	1722	US-10-505-928-780	Sequence 780, App
8	66	5.5	836	US-10-509-131-35	Sequence 35, App
9	65	5.4	331	US-11-118-524-9	Sequence 9, App
10	65	5.4	969	US-10-505-928-94	Sequence 94, App
11	63.5	5.3	5738	US-10-505-928-150	Sequence 150, App
12	63	5.3	555	US-11-242-505A-45	Sequence 45, App
13	62	5.2	314	US-10-538-066-366	Sequence 366, App
14	61.5	5.1	673	US-11-101-316-16	Sequence 16, App
15	61.5	5.1	799	US-10-505-928-716	Sequence 716, App
16	61	5.1	489	US-10-196-749-116	Sequence 116, App
17	61	5.1	442	US-10-505-928-259	Sequence 259, App
18	60.5	5.1	240	US-10-511-937-2977	Sequence 2977, App
19	60.5	5.1	459	US-10-370-959-60	Sequence 60, App
20	60.5	5.1	575	US-10-511-937-2625	Sequence 2625, App
21	60.5	5.1	999	US-10-196-749-434	Sequence 434, App
22	60.5	5.1	1821	US-10-505-928-451	Sequence 451, App
23	60	5.0	374	US-10-505-928-46	Sequence 46, App
24	60	5.0	382	US-10-196-749-20	Sequence 20, App
25	60	5.0	492	US-10-505-928-642	Sequence 642, App

26	59.5	5.0	738	US-10-511-937-2418	Sequence 2418, App
27	59	4.9	484	US-10-196-749-280	Sequence 280, App
28	59	4.9	484	US-11-101-316-78	Sequence 78, App
29	59	4.9	972	US-11-255-147-8	Sequence 8, App
30	59	4.9	1170	US-10-511-937-3007	Sequence 3007, App
31	58.5	4.9	769	US-11-246-999-67	Sequence 67, App
32	57.5	4.8	780	US-10-505-928-102	Sequence 102, App
33	57.5	4.8	481	US-11-249-111-77	Sequence 77, App
34	57	4.8	706	US-10-511-937-2450	Sequence 2450, App
35	57	4.8	728	US-11-311-555-18	Sequence 18, App
36	57	4.8	728	US-11-311-555-18	Sequence 18, App
37	57	4.8	2202	US-10-488-015-12	Sequence 12, App
38	57	4.8	4834	US-10-505-928-827	Sequence 827, App
39	56.5	4.7	264	US-10-196-749-12	Sequence 12, App
40	56	4.7	542	US-11-261-673-2	Sequence 2, App
41	55.5	4.6	638	US-11-302-678-47	Sequence 47, App
42	55.5	4.6	988	US-11-312-958-18	Sequence 18, App
43	55.5	4.6	2026	US-10-505-928-831	Sequence 831, App
44	55	4.6	408	US-10-505-928-455	Sequence 455, App
45	55	4.6	432	US-10-196-749-330	Sequence 330, App

ALIGNMENTS

RESULT 1
US-10-511-937-2425
Sequence 2425, Application US/10511937
Publication No. US2006008836A1
GENERAL INFORMATION:
APPLICANT: EXPRESSION DIAGNOSTICS, INC.
APPLICANT: Wohlgemuth, Jay
APPLICANT: Fry, Kirk
APPLICANT: Woodward, Robert
APPLICANT: Ly, Ngoc
APPLICANT: Prentice, James
APPLICANT: Morris, Macdonald
APPLICANT: Rosenberg, Steven
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
FILE REFERENCE: 506612000104
CURRENT APPLICATION NUMBER: US/10/511,937
CURRENT FILING DATE: 2004-10-19
PRIOR APPLICATION NUMBER: PCT/US2003/012946
PRIOR FILING DATE: 2003-04-24
PRIOR APPLICATION NUMBER: US 10/131,831
PRIOR FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: US 10/325,899
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 3117
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2425
LENGTH: 1006
TYPE: PRT
ORGANISM: Homo sapiens
US-10-511-937-2425

Query Match 6.0%; Score 71.5; DB 6; Length 1006;

Best Local Similarity 21.1%; Pred. No. 4.4;

Matches 56; Conservative 30; Mismatches 79; Indels 101; Gaps 14;

QY	4	QPLRHRSRCATPPRGDF-----CCGTERAIDQASFTTSMEMDQVVGKSSPLAP	52
DB	291	QPARGDKACQACPRGLYKSSAGNAPCSPPARSHAFNFA-----APVCP	334
QY	53	AGIG-----AEEPAAGPOLPS-----WLPERCAYVQCAQCHAVLADSVHAWDLER	99
DB	335	CLEGFRRASDDPEAPCTGP--PSAPQBLMFE-----VQSSALMTKRLR	378
QY	100	SIQA---VFESVTNNVLEAPFLVIGSLKSTYNNLLFCGSCGIPVGFH---LST	151
DB	379	ELGGRODLIF-----NVCKE-----CEGRORPASGGGCTCHRCRDEVFDFPRORGLTES	428

```

OY      152  AAAAAAAGHCTSSDKMVCYLKTKRAIVMAGEMDIONVPLEKIAELKEKIVLTHNRUK 211
Db      429  RVVGGGLAAH-----VPYILEVQAVGVSELS-----PDPQAAAIN---VSTSHVFP 473
OY      212  SLMKILSEVT-----PDQS 225
Db      474  SAVPVHGVSRASNSITVSWPQPDQT 499

RESULT 2
US-11-312-958-44
; Sequence 44, Application US/11312958
; Publication No. US20060100152A1
GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Rosenfeld, Julie Beth
; APPLICANT: Silos-Santiago, Immaculada
TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
TITLE OF INVENTION: 27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
TITLE OF INVENTION: 13424 MOLECULES
FILE REFERENCE: MP102-027P1RNMNMIM
CURRENT APPLICATION NUMBER: US/11/312, 958
CURRENT FILING DATE: 2005-12-20
PRIOR APPLICATION NUMBER: US/10/369, 022
PRIOR FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US 60/360,495
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/370,121
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/373,010
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/373,908
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/377,717
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/379,949
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: US 60/382,409
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US 60/385,280
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 60/386,879
PRIOR FILING DATE: 2002-06-06
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-312-958-44

Query Match      5.6%; Score 67; DB 7; Length 388;
Best Local Similarity 25.0%; Pred. No. 3.7;
Matches 29; Conservative 15; Mismatches 42; Indels 30; Gaps 4;

OY      49  PLGPGAGLGEPRAGQLPSWLOPERCAVPGCAQCAHAVLADSV---HLMADLSRSLGAVY 105
Db      59  PLPPPGGL-----TPRLHARQLVYAACVCFVPMAGVGVGLAHSLA----- 101

OY      106  FSRVTNNVVAEAPFLVIGIEGSLKGYTNLLFC---GSCGIPVGFHDYSTHAALAL 158
Db      102  -----IMTDAHLADVGSMMGSLFSLMLSTRPATRTMTPEMHNSSETLIGALASY 150

RESULT 3
US-10-370-959-152
; Sequence 152, Application US/10370959
; Publication No. US20060088907A1
; GENERAL INFORMATION:

```

```

APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark W.
APPLICANT: Kapeller-Libermann, Roana
APPLICANT: Macbeth, Kyle J.
APPLICANT: Hunter, John J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Bandaru, Rajasekhar
APPLICANT: Tezi, Fong-Ying
TITLE OF INVENTION: NOVEL 11237, 18480, 2245, 16228, 7677,
TITLE OF INVENTION: 26320, 46619, 33366, 16636, 46867, 22617, 55562, 39228,
TITLE OF INVENTION: 62088, 46745, 23155, 21657, 42755, 32229, 22325, 46863, AND
TITLE OF INVENTION: 32252 MOLECULES AND USES THEREFOR
FILE REFERENCE: MP103-017ONMIM
CURRENT APPLICATION NUMBER: US/10/370, 959
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: US 09/910, 150
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/219, 028
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: US 10/251, 507
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: US 09/715, 479
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: US 60/218, 053
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: US 09/644, 929
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 60/212, 439
PRIOR FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: US 09/892, 870
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: US 60/214, 174
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: US 09/775, 117
PRIOR FILING DATE: 2001-02-01
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PASTESEQ for Windows Version 4.0
SEQ ID NO 152
LENGTH: 672
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-370-959-152

Query Match 5.6%; Score 67; DB 6; Length 672;
Best Local Similarity 20.6%; Pred. No. 8.1;
Matches 36; Conservative 31; Mismatches 60; Indels 48; Gaps 7;

QY 95 MDLSRLGAVFESVNTNNVLEAPFLVIGBSLKGSTN-----LLFGSC 140
DB 66 WKFS-----GIVFSRVYBEVDTSKGIDVPEWFGSRNLVANELLRKENDVALYIARE 121
QY 141 G-----IPVGFHLVSTHAL--AALRGHFCLSDKRWCVYLTK-----AIVN 181
DB 122 GKKEIVKVTPEELAQEVALPAAARKVKGVKKGDAVVGILPSEHVAEAMLAASIGAIWS 181
QY 182 ASENDIONVPLSEKIAELKEKIVLT-----HNLKSLMKIILSEVTPDQK 226
DB 182 STSPDFGVNGVLDRFSQIQPKLIFSVEAVVYNGKEHNHMERLQGVKGL-PDLKK 235

RESULT 4
US-11-181-115-5
; Sequence 5, Application US/11181115
; Publication No. US20060088511A1
; GENERAL INFORMATION:
; APPLICANT: Dana Farber Cancer Center
; TITLE OF INVENTION: Cancer Therapy Sensitizer
; FILE REFERENCE: 7032/2072
; CURRENT APPLICATION NUMBER: US/11/181, 115
; CURRENT FILING DATE: 2005-07-14
; PRIOR APPLICATION NUMBER: PCT/US04/000901

```

; PRIOR FILING DATE: 2004-01-14
 ; PRIOR APPLICATION NUMBER: US 60/440,009
 ; PRIOR FILING DATE: 2003-01-14
 ; NUMBER OF SEQ ID NOS: 81
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 5
 ; LENGTH: 345
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-181-115-5

Query Match 5.6%; Score 66.5; DB 7; Length 345;
 Best Local Similarity 25.0%; Pred. No. 3.6;
 Matches 16; Conservative 8; Mismatches 15; Indels 25; Gaps 2;

QY 47 SSPLGPAAGAEPPAAGP-----QLPSWLOPERCAV-----FOCA 81
 DB 278 SAMPAPPAVGGGBCAPCCCLPQETVAVVQCDGCVWFHVACVCGSIQAAREADPRCP 337

QY 82 QCHA 85
 DB 338 GCRA 341

RESULT 5
 US-11-181-115-8
 ; Sequence 8, Application US/11181115
 ; Publication No. US20060088511A1
 ; GENERAL INFORMATION:

; APPLICANT: Dana Farber Cancer Center
 ; TITLE OF INVENTION: Cancer Therapy Sensitizer
 ; FILE REFERENCE: 7032/2072
 ; CURRENT APPLICATION NUMBER: US/11/181,115
 ; PRIOR FILING DATE: 2005-07-14
 ; PRIOR APPLICATION NUMBER: PCT/US04/000901
 ; PRIOR FILING DATE: 2004-01-14
 ; PRIOR APPLICATION NUMBER: US 60/440,009
 ; PRIOR FILING DATE: 2003-01-14
 ; NUMBER OF SEQ ID NOS: 81
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 8
 ; LENGTH: 345
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-181-115-8

Query Match 5.6%; Score 66.5; DB 7; Length 345;
 Best Local Similarity 25.0%; Pred. No. 3.6;
 Matches 16; Conservative 8; Mismatches 15; Indels 25; Gaps 2;

QY 47 SSPLGPAAGAEPPAAGP-----QLPSWLOPERCAV-----FOCA 81
 DB 278 SAMPAPPAVGGGBCAPCCCLPQETVAVVQCDGCVWFHVACVCGSIQAAREADPRCP 337

QY 82 QCHA 85
 DB 338 GCRA 341

RESULT 6
 US-11-181-115-12
 ; Sequence 12, Application US/11181115
 ; Publication No. US20060088511A1
 ; GENERAL INFORMATION:

; APPLICANT: Dana Farber Cancer Center
 ; TITLE OF INVENTION: Cancer Therapy Sensitizer
 ; FILE REFERENCE: 7032/2072
 ; CURRENT APPLICATION NUMBER: US/11/181,115
 ; PRIOR FILING DATE: 2005-07-14
 ; PRIOR APPLICATION NUMBER: PCT/US04/000901
 ; PRIOR FILING DATE: 2004-01-14
 ; PRIOR APPLICATION NUMBER: US 60/440,009
 ; PRIOR FILING DATE: 2003-01-14

; NUMBER OF SEQ ID NOS: 81
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 12
 ; LENGTH: 345
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-181-115-12

Query Match 5.6%; Score 66.5; DB 7; Length 345;
 Best Local Similarity 25.0%; Pred. No. 3.6;
 Matches 16; Conservative 8; Mismatches 15; Indels 25; Gaps 2;

QY 47 SSPLGPAAGAEPPAAGP-----QLPSWLOPERCAV-----FOCA 81
 DB 278 SAMPAPPAVGGGBCAPCCCLPQETVAVVQCDGCVWFHVACVCGSIQAAREADPRCP 337

QY 82 QCHA 85
 DB 338 GCRA 341

RESULT 7
 US-10-505-928-780
 ; Sequence 780, Application US/10505928
 ; Publication No. US20060088532A1
 ; GENERAL INFORMATION:

; APPLICANT: Ludwig Institute for Cancer Research et al.
 ; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
 ; FILE REFERENCE: 28967/39178
 ; CURRENT APPLICATION NUMBER: US/10/505,928
 ; PRIOR FILING DATE: 2004-08-27
 ; PRIOR APPLICATION NUMBER: US 60/363,019
 ; PRIOR FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 866
 ; SOFTWARE: PatentIn 3.2
 ; SEQ ID NO 780
 ; LENGTH: 1722
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-505-928-780

Query Match 5.6%; Score 66.5; DB 6; Length 1722;
 Best Local Similarity 22.5%; Pred. No. 35;
 Matches 18; Conservative 6; Mismatches 15; Indels 41; Gaps 2;

QY 35 TSMEMDTQVVGSSPLGPAAGAEPPAAGFQPSWLOPERC-----75
 DB 784 TKLEWCOIPIKGR-----FKTDPWYNPDRAGIHGPPLIIEGSEYFVAD 828

QY 76 -----AVFQCAQCHAVLA 88
 DB 829 LHLNREAVLYCCASHSFILA 848

RESULT 8
 US-10-509-131-35
 ; Sequence 35, Application US/10509131
 ; Publication No. US20060089493A1
 ; GENERAL INFORMATION:

; APPLICANT: Biogen Idec Inc.
 ; APPLICANT: McLachlan, Karen
 ; TITLE OF INVENTION: NOVEL GENE TARGETS AND LIGANDS THAT BIND THERETO FOR TREATMENT OF
 ; FILE REFERENCE: 037003-0302886
 ; CURRENT APPLICATION NUMBER: US/10/509,131
 ; PRIOR FILING DATE: 2004-09-28
 ; PRIOR APPLICATION NUMBER: US 60/367,727
 ; PRIOR FILING DATE: 2002-03-28
 ; PRIOR APPLICATION NUMBER: US 60/381,328
 ; PRIOR FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: US 60/386,747
 ; PRIOR FILING DATE: 2002-06-10

QY 135 -----LFC--GSGCIPVGFHLYSTHA 153
DB 3263 CSQCVCERKBLCCPGCGCPFLPCGMSAMSWA 3293

RESULT 12
US-11-242-505A-45
; Sequence 45, Application US/11242505A
; Publication No. US20060099656A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Healy, Aileen
; TITLE OF INVENTION: Methods and Compositions for Treating
; TITLE OF INVENTION: Hematological Disorders Using 232, 2059, 10630, 12848, 13875,
; FILE REFERENCE: MPI2001-288P1C10N1M
; CURRENT APPLICATION NUMBER: US/11/242, 505A
; PRIOR FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US 10/290, 078
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: US 60/347, 949
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/320, 351
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341, 606
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-242-505A-45

Query Match
Best Local Similarity 26.1%; Pred. No. 18;
Matches 18; Conservative 6; Mismatches 27; Indels 18; Gaps 3;

QY 37 MEMPTQVVKSSP-----LCPAGLGAEPFPAQQLSWIQPER-CAVFCQAQ 82
DB 264 IHMQASARPPGPIEBRGEGDHNHDTGEGGGEPRGQ----RWDEQMSVVCED 319

QY 83 CHAVIADSV 91
DB 320 CELIPADHV 328

RESULT 13
US-10-538-066-366
; Sequence 366, Application US/10538066
; Publication No. US20060094649A1
; GENERAL INFORMATION:
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: HLA-A1, -A2, -A3, -A24, -B7, and -B44 Tumor Associated Antigen
; FILE REFERENCE: 2060 015PC06
; CURRENT APPLICATION NUMBER: US/10/538, 066
; PRIOR FILING DATE: 2005-06-09
; PRIOR APPLICATION NUMBER: US 60/432, 017
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 767
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 366
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-538-066-366

Query Match
Best Local Similarity 5.2%; Score 62; DB 6; Length 314;
Matches 40; Conservative 20; Mismatches 79; Indels 42; Gaps 7;

QY 5 PLHRSRCATPPRG-----DFCGTERAIDQASFTTSMWMDTQVVKSSPLGPAGLG 56

DB 2 PLEGRSGHCRPEBELARGELGVGAQAPRTEGEAASSSTLVEVTLGVPAAESPD 61
QY 57 AEPPAGPOLPS-----WIO-----PERCAVQCQACHAVIADSVHLA 94
DB 62 PQSPQASSLPTTNNVPLWMSQSYEDSSNOEBEGPSTPDLSEFOALSRKV-ALVHFL 120
QY 95 WDLRSIGAVFESVTNNVLEAPFLVIGESLKGSTYNLLFCGSCG1-----PVGFHL 148
DB 121 LKTRAREPVTKAMLSVGNWQYFPFVIFSKASSLQVf-----GIELMEVPDPIG-HL 175

QY 149 Y 149
DB 176 Y 176

RESULT 14
US-11-101-316-16
; Sequence 16, Application US/11101316
; Publication No. US20060099657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; FILE REFERENCE: P3230R1C17C1
; CURRENT APPLICATION NUMBER: US/11/101, 316
; PRIOR FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 16
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-16

Query Match
Best Local Similarity 5.1%; Score 61.5; DB 7; Length 673;
Matches 31; Conservative 7; Mismatches 43; Indels 27; Gaps 7;

QY 10 SRCATPPRGDFCGTERAIDQASP-----TTSMEWDTQVVK-----SSPLGPAGIAG 58
DB 326 TRCHFPPKN--AGRLLELDYADPGCAPTTTATVPTTRVVRPPLALSSSLAPTWLSPT 383

QY 59 EPAA-GPOLPSWL-----QPERCAVFOC--AOCHAVIADSVHLA 94
DB 384 APATEAPSPSTAPPTVGPVPQDDCPPTCLNGTCH--LGTTHHLA 429

RESULT 15
US-10-505-928-716
; Sequence 716, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505, 928
; PRIOR FILING DATE: 2004-08-27

